THE GENETIC ARCHITECTURE OF QUANTITATIVE DISEASE RESISTANCE IN MAIZE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jesse Abner Poland

February 2010

THE GENETIC ARCHITECTURE OF QUANTITATIVE DISEASE RESISTANCE IN MAIZE

Jesse Abner Poland, Ph.D.

Cornell University 2010

Several large scale quantitative genetic studies were conducted to better understand the genetic basis for quantitative disease resistance (QDR) in plants. The focus of these studies was the economically important disease of maize (*Zea mays* L. ssp. mays), northern leaf blight (NLB, caused by *Setosphaeria turcica* L. anamorph *Exserohilum turcicum*). The maize nested association mapping (NAM) population, a reference design population consisting of 4,630 recombinant inbred lines, was evaluated over three environments for quantitative resistance to NLB, giving highly heritable resistance phenotypes. Over 200 resistance alleles at 30 different quantitative trait loci (QTL) for disease resistance were identified. Genome-wide nested association mapping for NLB resistance identified genes at six of the QTL that have been associated with disease resistance including three receptor-like kinases, two ethylene response factors, and one *Mlo*-like gene. Further insight on QDR, with a focus on multiple disease resistance (MDR), was gained by jointly analyzing independent data on NAM for resistance to southern leaf blight (SLB), gray leaf spot (GLS) and NLB. To examine the possibility of MDR genes, the estimated allele effects from each founder inbred were compared at loci were QTL for two or more diseases co-localized. At seven loci, positively correlated allele effects provided evidence for MDR genes. Analysis of the NAM population suggested that resistance to the three

diseases studied here is largely due to the accumulation of disease-specific genes and, to a limited extent, pleiotropic genes that condition MDR.

A final study was conducted to determine the effect of variability in visual disease rating on mapping disease QTL by assessing the effects of scorer variability and rating scales on mapping QTL for NLB in a single recombinant inbred line population from NAM. Stepwise general linear model selection (GLM) and inclusive composite interval mapping (ICIM) were used for QTL mapping. For both GLM and ICIM the same QTL were largely found across scorers, though some QTL were only identified by some scorers. Strikingly, the magnitudes of estimated allele effects from different scorers at identified QRL were drastically different, sometime by as much as three fold.

The studies conducted here advance the understanding of QDR in plants and lay groundwork for identifying the genes responsible for resistance to NLB in maize. A greater understanding of QDR will assist in the development of durable resistant crop cultivars, improving food security and safety.

BIOGRAPHICAL SKETCH

Jesse Poland was born in rural Kansas, the son of George and Connie Poland, and grew up on the family farm. It was here that he learned the value of hard work and gained a love for agriculture. He has many fond memories and learned many valuable lessons from time spent working with his father on the farm. Returning to this part of Kansas, where his father and uncle continue to manage the family farm, always feels like home.

Jesse began his (long) university career at Kansas State University studying agronomy. It was in the course of his studies that an interest in genetics was sparked. After completing an introductory genetics class, he managed to convince one of the agronomy faculty, Gina Brown-Guedira, to take him on as an undergraduate research assistant. It was during this time that Jesse gained valuable research experience and realized the importance of plant breeding around the world. It was also during this time that he became a Christian and developed many close friendships with his FarmHouse Fraternity brothers. Also at Kansas State, Jesse met his wife, Ashley and one week after graduation they were married.

Jesse continued at Kansas State working with Scot Hulbert on the maize resistance gene *Rp1*. His studies focused on characterization of unequal recombination in this large gene family. In 2004, he completed a Master's of Science degree in Plant Pathology. During the spring of 2005, Jesse and his wife traveled to Thailand and Vietnam, working on an agriculture development project.

In the fall of 2005, Jesse began graduate work in Plant Breeding and Genetics at Cornell University. He has continued in this degree program gaining many valuable life lessons, learning more statistics than he ever thought he would need, and gaining experience in international agriculture. It

was primarily for this international experience that Jesse was attracted to Cornell. His research has focused on understanding the genetics of quantitative disease resistance in maize, with the goal of making larger contributions to breeding durably resistant plant cultivars.

The most recent milestone of Jesse's life was the birth of his son, Thatcher, on August 4, 2009 (during the peak of the maize field season).

For my wife, Ashley,

whose encouragement and sacrifice made this happen

ACKNOWLEDGMENTS

I would like to give special acknowledgement to my advisor Rebecca Nelson. It is always a little risky when you commit to a new student so I thank you for taking the gamble.

I appreciate those who got me started in life and research. Thank you to my parents, George and Connie Poland, for their support and encouragement through school and for helping me develop a strong work ethic. I would also like to acknowledge my past advisors. First, Gina Brown-Guedira, for giving me the opportunity to get started in research. I would also like to acknowledge Scot Hulbert for his guidance in my Master's program.

My graduate committee here at Cornell has been excellent, giving valuable advice on every aspect of my research. I greatly appreciate their input and help, despite the great difficulties of getting them all in the same room at the same time. Thank you, Margaret Smith, Gary Bergstrom, and Edward Buckler.

I appreciated very valuable input from Randall Wisser, Peter J. Balint-Kurti, and Richard C. Pratt for Chapter 1.

For Chapters 2 and 3, I greatly appreciate help from past and present members of the R. Nelson lab for assistance in putting out field trials for NAM; Oliver Ott, Ellie Walsh, Chia-lin Chung, Lin-Si Hseih, Judith Kolkman, Joy Longfellow, Tiffany Jamann, Santiago Mideros, Kristen Kennedy, Kerri Lyons, Cassilyn Schweighofer, Ariel Fialko, Sara Heins. Oliver Ott and Ellie Walsh greatly contributed to phenotyping incubation period in 2007. I appreciate

intellectual support from members of the Buckler Lab. In particular, Nick Lepak for assistance with seed sources and Dallas Kroon for assistance with developing a database and bar-coding system. Help from Jim Holland on ASReml mixed models was greatly appreciated. I also had many valuable discussions with Peter Bradbury, receiving much valuable advice and programs.

For chapter 3, the phenotyping of southern leaf blight was conducted by the James Holland and Peter Balint-Kurti groups in North Carolina State University. This was Kristen Kump thesis project and I have enjoyed working with her. Randall Wisser contributed to southern leaf blight phenotyping and review of the paper. Jacqueline Benson, in collaboration with Erik Stromberg at Virginia Tech. conducted all of the gray leaf spot phenotyping. Peter Bradbury and Edward Buckler gave excellent technical support on the data analysis.

The evaluation of scorer variability in chapter 4 was made possible by all of the scorers who participated in the study. Thank you to Peter Balint-Kurti, Chialin Chung, Sarah Collier, Ariel Fialko, Elliot Heffner, Linsi Hsieh, Tiffany Jamann, Kristen Kennedy, Judy Kolkman, Kerri Lyons, Patricia Manosalva, Alexandre Mello, Santiago Mideros, Oliver Ott, Cassilyn Schweighofer, Sumita Sen, Feng Tian, Nengyi Zhang, Zhiwu Zhang and Paola Zuluaga. Hiroyoshi Iwata developed the genomic selection program and Elliot Heffner assisted in running the program.

For chapter 5 I appreciated the help of Cassilyn Schweighofer, Ariel Fialko, Sarah Heins with tissue collection, dna extraction and genotyping. Chialin

Chung had previously characterized the lines used for fine-mapping and provided extremely valuable input.

Oliver Ott contributed to the development of the lines and phenotyping for southern leaf blight described in chapter 6. Peter Balint-Kurti hosted the field trails southern leaf blight. Jacqueline Benson and Erik Stromberg evaluated gray leaf spot on the populations.

I would also like to acknowledge everyone in the Plant Breeding Department and the Plant Pathology Department. This has been an excellent graduate experience. My studies have been greatly enriched through the interaction with my fellow students and the outstanding faculty of our departments.

I am grateful to the Department of Plant Breeding and Genetics, The McKnight Foundation, the Generation Challenge Program, and the Bill and Melinda Gates Foundation for financial support. The Cornell Graduate School provided support for travel to several conferences.

Finally, I am very thankful for the support and encouragement that my wife, Ashley, has given. I give thanks to the Lord for giving me a very wonderful family to be with me through my studies and for giving me the opportunity to complete my graduate degree.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Shades of gray: the world of quantitative disease resistance[1]

ABSTRACT

A thorough understanding of quantitative disease resistance (QDR) would contribute to the design and deployment of durably resistant crop cultivars. The molecular mechanisms that control QDR remain poorly understood, however, due largely to the incomplete and inconsistent nature of the resistance phenotype, which is usually conditioned by many loci of small effect. Here we discuss recent advances in research on QDR. Based on inferences from analyses of the defense response and from the few isolated QDR genes, we suggest several plausible hypotheses for a range of mechanisms underlying QDR. We propose that a new generation of genetic resources, complemented by careful phenotypic analysis, will produce a deeper understanding of plant defense and lead to more effective utilization of natural resistance alleles.

The two worlds of disease resistance

Two general categories of disease resistance have long been recognized in plants (e.g. Ref. (van der Plank 1968)) : (i) complete resistance conditioned by a single gene and (ii) incomplete resistance conditioned by multiple genes of partial effect. In their extreme forms, these types of

---

[1] Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ. 2009. Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science* 14: 21-9

resistance are clear and easily distinguished. A variety of terms have been used to refer to this perceived dichotomy including: horizontal vs. vertical, complete vs. incomplete, major-gene vs. minor-gene, and narrow-spectrum vs. broad-spectrum. This diversity of terms reflects the range of interests and assumptions made by the respective authors, but adds an element of confusion to the literature as some terms are used in different ways by different authors. Here, we use the terms “qualitative resistance” and “quantitative resistance” (QDR; see glossary) to refer to the respective phenomena. We use the term “R-genes” to refer to genes that confer qualitative effects in a gene-for-gene manner, and “QRL” (quantitative resistance loci; see glossary) (Young 1996) to refer to loci or genes that confer quantitative resistance. Although the phenomena of qualitative and quantitative resistance can be considered different, there is a great deal of gray area between the extremes, suggesting that it might be useful to reexamine the concepts in light of emerging evidence on mechanisms of resistance. As we discuss, several authors have questioned whether loci controlling the two types of resistance are distinct (Kamoun *et al.* 1999), suggesting that quantitative and qualitative resistance are conditioned by the same genetic mechanisms.

There is substantial support for this hypothesis in some cases, as well as evidence that alternative mechanisms also underlie quantitative resistance (e.g., Refs. (Andaya and Ronald 2003; Fu *et al.* 2009; Fukuoka *et al.* 2009; Krattinger *et al.* 2009)). Several credible hypotheses can be proposed, and it is likely that a diversity of mechanisms will be implicated, some overlapping with qualitative resistance (e.g. specific recognition of pathogen effectors or their targets) and innate immunity (e.g., relatively non-specific recognition, such as recognition of broadly conserved pathogen features; also known as

basal resistance). Other plausible mechanisms to explain quantitative resistance would include the detection and mitigation of infection-related damage, the modulation and transduction of defense signals, and mechanisms of direct defense (e.g., establishment or reinforcement of defensive structures, antimicrobial secondary metabolites, or detoxification of pathogen-derived toxins). The mechanisms involved in conditioning QDR are likely to have implications for resistance spectra associated with specific QRLs, as well as for their durability. Mechanisms of specific recognition might eventually be overcome as a result of pathogen evolution, whereas non-specific defense mechanisms could provide resistance that is relatively broad in spectrum and robust in the face of pathogen evolution.

Although various authors have speculated on the types of genes that might underlie QRLs, the evidence to date has often relied on co-localization of QRLs and genes in low-resolution mapping studies (Wang *et al.* 2001; Liu *et al.* 2004). In this situation, limited inference can be made because hundreds of genes are typically located in genomic regions defined by QRLs. In some cases, however, positional cloning has been achieved, and it is expected that more QRLs will be isolated. Here, we synthesize and interpret the literature pertaining to the mechanisms of quantitative resistance in plants, focusing primarily on fungal pathosystems with some reference to bacterial systems. Viral systems are not considered because the pathogenic strategies of viruses are so different from those of fungi and bacteria. We provide a brief overview of qualitative and quantitative resistance, assess recent developments in understanding quantitative resistance in crop and model species, place these findings in the context of the broader understanding of plant defense, and suggest ways in which greater synergies could be achieved.

The limitations of qualitative resistance

R-genes typically provide high levels of resistance and are relatively easy to manipulate, both in basic research and applied breeding programs. These genes are important in many systems, but their utility varies among pathosystems (see glossary) and among genes within a pathosystem. The primary limitations of R-genes for crop protection are (i) a lack of durability in some systems (primarily with respect to pathogens that have high evolutionary potential (McDonald and Linde 2002)) and (ii) a lack of availability in others (primarily necrotrophic systems).

The ephemeral nature of R-gene mediated resistance is highlighted by a recent outbreak of a new strain of wheat stem rust (caused by *Puccinia graminis* race Ug99) that is virulent on cultivars carrying widely deployed resistance genes (Stokstad 2007). The deployment and breakdown of R-genes (see glossary) for diseases such as wheat rust has been a frustrating battle for plant breeders, pathologists and farmers, and more durable resistance is needed (Ayliffe *et al.* 2008). For potato late blight, caused by *Phytophthora infestans*, lines carrying most R-genes have been rapidly overcome by virulent populations. In addition to subjecting pathogen populations to high levels of selection pressure, the presence of R-genes shifts the trait distribution in such a way that underlying quantitative trait loci cannot be detected. This phenomenon has led some breeding programs to eliminate R-genes intentionally so that QDR could be more effectively assessed and advanced (Landeo *et al.* 1995).

Several recent reviews have discussed current insights on R-gene-mediated resistance (Nimchuk *et al.* 2003; Jones and Dangl 2006; Bent and Mackey 2007). Many R-genes have been isolated and characterized, and the downstream responses that they trigger are increasingly well understood. R-

genes typically encode proteins that recognize pathogen effectors or modifications of plant proteins that are targets of those effectors (Nimchuk *et al.* 2003). Among the six classes of R-genes, the most common class contains characteristic nucleotide binding and leucine rich repeat (NB-LRR; see glossary) amino acid sequence motifs involved in recognition and related functions. However, other mechanisms have been associated with qualitative resistance, including detoxification of fungal toxins (e.g., *Hm1* and *Hm2* in maize (Johal and Briggs 1992)), modulation of the defense response (e.g., *mlo* in barley (Büschges *et al.* 1997)) and transcriptional regulation (e.g., *Bs3* in pepper (Romer *et al.* 2007)).

R-genes mediate a highly effective defense response to invasion by biotrophic pathogens, which usually involves a hypersensitive response (HR) in which tissue immediately adjacent to the site of pathogen ingress undergoes rapid programmed cell death. However, the same response can increase susceptibility to necrotrophic pathogens. For a pathogen that thrives on dead host tissue, exploiting programmed cell death in the host is a perfect method for acquiring nutrients. This is illustrated in two studies showing that necrotrophic pathogens can produce host-specific toxins that activate R-gene mediated defense responses resulting in host cell death ((Lorang *et al.* 2007; Nagy *et al.* 2007), J. Bennetzen personal communication). Thus, it is not surprising that R-gene-mediated defense to necrotrophic pathogens is rare. The few known naturally-occurring qualitative resistance genes for true necrotrophs encode detoxification enzymes rather than genes that mediate the HR (e.g., Ref. (Johal and Briggs 1992; Brandwagt *et al.* 2002; Sindhu *et al.* 2008)). However, quantitative genetic resistance to necrotrophic pathogens is typically available and can be effective (Figure 1, e.g., Ref. (Balint-Kurti *et al.* 2008)).

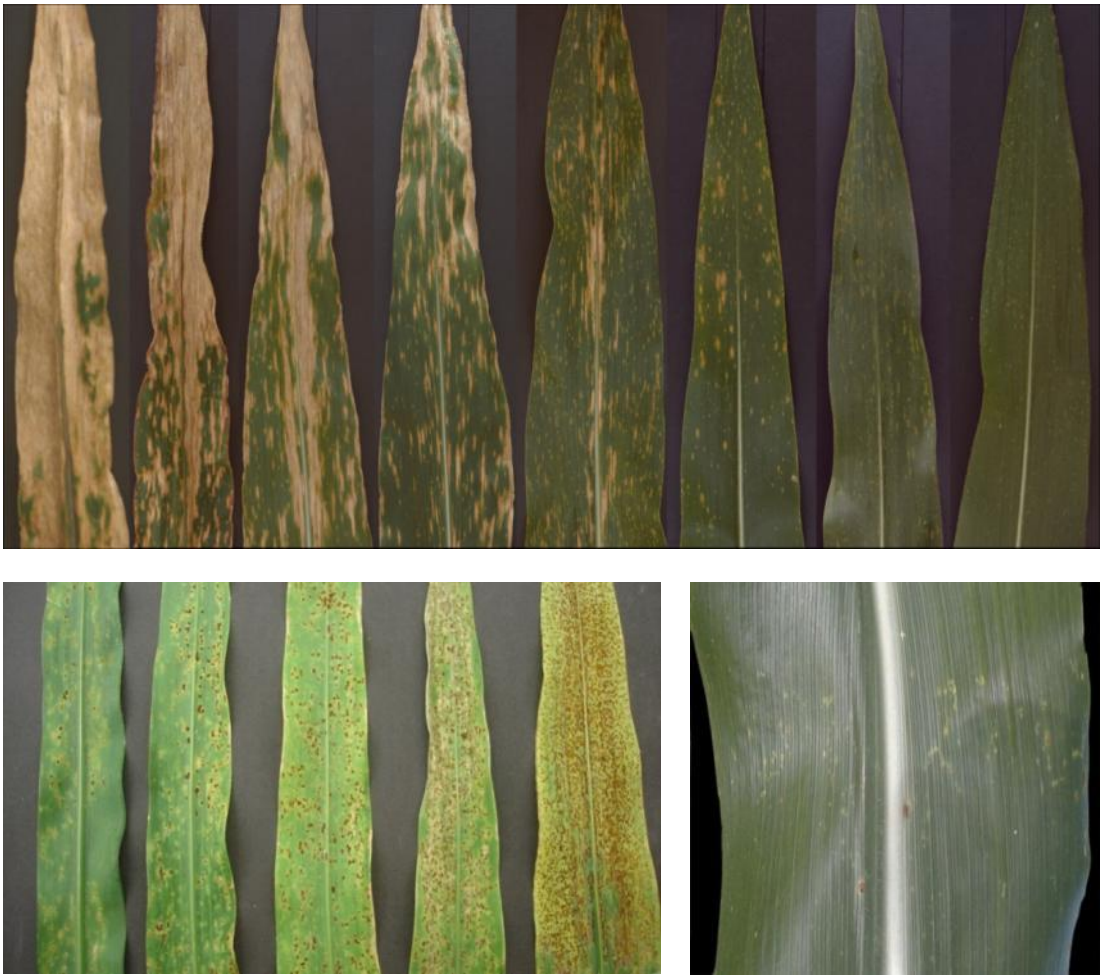


Figure 1.1. Examples of quantitative disease resistance in maize

1.1a. Quantitative variation for resistance to *Cochliobolus heterostrophus*, the causal agent of southern leaf blight, in a segregating maize population. Each leaf is representative of a recombinant inbred line derived from a cross between maize inbred lines B73 and De811. The range of differences in levels of foliar blight indicates that there is large variation in QDR for this pathogen. This quantitative resistance can be highly effective, even though R-genes are not present for resistance to *C. heterostrophus*. (*Reproduced with permission from Chi-ren Shyu, Jason Green*)

1.1b. Range in quantitative resistance to *Puccinia sorghi* in a segregating maize population. Contrasting to
1.1c showing a hypersensitive resistance reaction from the R-gene *Rp1*. Although the resistance provided by R-genes is very effective, it is often subject to breakdown resulting from pathogen evolution. (*Photo 1b: Jesse Poland, Photo 1c reproduced with permission from Jerald Pataky*)

What is quantitative resistance, and why should we care about it?

QRL have been identified in most (and perhaps all) plant pathosystems. We identified from the literature 25 QRL mapping studies for rice diseases, 43 for maize diseases and 13 for *Arabidopsis* (http://www.plantpath.cornell.edu/Labs/Nelson_R/TIPS2008_QRLcitations.html). The large number of studies in crop species reflects the importance of QDR in agricultural production. QDR has been reported in several *Arabidopsis* pathosystems, including *Pseudomonas syringae* (Kover and Schaal 2002), *Erysiphe cichoracearum* (Wilson *et al.* 2001) and *Botrytis cinerea* (Denby *et al.* 2004). Variation in QDR is important for crop improvement and can be selected, often leading to high levels of phenotypic resistance (Singh *et al.* 2005). For some plant diseases, such as rice blast and bacterial blight, resistance breeding relies on a combination of quantitative and qualitative forms of resistance. For other diseases, including those caused by necrotrophic pathogens, QDR is the most important or only form of resistance available (Figure 1).

Observations of the performance of crop cultivars with different types of resistance has led to the conclusion that QDR tends to be more durable than does typical R-gene-mediated resistance (Parlevliet 2002). Whereas R-genes

can be rapidly overcome by strong selection for compatible pathogen variants, resistance breakdown (see glossary) is considered to be less of a problem with QRLs, because of their smaller effects (leading to lower selection pressure on the pathogen) and/or presumed broader specificity (the latter being widely assumed but little documented). Because QDR is controlled by multiple genes with partial and inconsistent effects, pathogen variants that overcome QRLs gain only a marginal advantage.

Quantitative disease resistance in context of the current model of plant-pathogen interactions

Jones and Dangl (Jones and Dangl) recently summarized the complex interplay between (biotrophic) pathogen attack and host defense as a multi-phase 'zig-zag' process in which the evolutionary arms race between host and pathogen results in an oscillation between compatible (susceptible) and incompatible (resistant) states over time (see glossary). The host plant initially recognizes features common to many microbes, such as flagellin or chitin (microbial- or pathogen-associated molecular patterns, MAMPs) using pattern recognition receptors (see glossary). This recognition event then triggers the innate immune response (also known as host basal defenses), such as cell wall apposition, arresting further pathogen development (i.e., MAMP-triggered immunity). Successful pathogens evade basal defenses using effector proteins that disrupt the normal defense response. In turn, host plants have evolved NB-LRR proteins that recognize these pathogen effectors and mount heightened defense responses (i.e., effector triggered immunity). The loss or mutation of specific pathogen effector proteins enables avoidance of R-gene recognition and virulence. R-gene recognition and corresponding losses and mutations of effectors (avr genes) leads to the widely recognized "gene-for-

gene" interactions. While this type of interaction is typified by a strong resistance or susceptibility phenotype, modifier loci can affect the strength of R gene-mediated defense (e.g., Refs. (Hu *et al.* 1997; Bryan *et al.* 2004)). As there are multiple genes involved in the resistance pathway, natural functional mutations could introduce quantitative variation to several or all of the phases described in the zig-zag model, adding shades of gray to the extremes of complete resistance and susceptibility. Nonetheless, it is unlikely that this model accounts for all known forms of QDR. The gene-for-gene model has been primarily constructed based on observations with biotrophic pathogens, for which R-gene mediated recognition and defense is effective.

Mechanisms underlying QRLs: hypotheses, credibility, evidence and proof

As an increasingly broad range of microbial pathogenic strategies and a corresponding range of host defense strategies are recognized, a corresponding array of molecular mechanisms can be postulated as playing a role in QDR. We highlight several hypotheses below, and outline evidence pertaining to each. Even at this early stage it is clear that more than one hypothesis is likely to be valid and that no single hypothesis can fully explain the breadth of QDR. Future work will likely suggest additional hypotheses and mechanisms.

Hypothesis #1. Quantitative resistance is conditioned by genes regulating morphological and developmental phenotypes

Plant diseases develop within the spatiotemporal context of plant development, so it is reasonable to speculate that some QRL are based on genes that control plant architecture or development. It has been well

documented in many necrotrophic plant-pathogen systems that flowering time is strongly correlated with disease resistance, such that susceptibility is apparently enhanced after flowering (e.g., Ref. (Collins *et al.* 1999)). Other developmental-stage-specific QRL have been documented (Thompson and Bergquist 1984; Steffenson *et al.* 1996). Morphological traits, such as stomatal density and/or openness or the ability to repel water can have important effects on disease resistance (Bradley *et al.* 2003; Melotto *et al.* 2006) in addition to other aspects of plant morphology, such as plant height, leaf area and leaf angle (Albar *et al.* 1998; Zhu *et al.* 1999). Therefore, it is likely that genes affecting growth and development, as well as plant architecture, have pleiotropic effects on disease resistance.

Hypothesis #2. QRLs represent mutations or different alleles of genes involved in basal defense

Flagellin (the main constituent of bacterial flagella) and chitin (the main component of fungal cell walls) are two widely conserved pathogen features that enable plants to recognize broad pathogen groups. Work with the *Arabidopsis* gene *Fls2*, an LRR receptor–like kinase involved in the perception of flagellin, has established a mechanistic basis of flagellin recognition and basal resistance in plants (Zipfel *et al.* 2004). Mutations in *Fls2* appears to produce a phenotype of modest (quantitative) effects on disease severity and bacterial colonization when tested under natural conditions (Zipfel *et al.* 2004) and different alleles of *Fls2* have been found in several *Brassica* species (Dunning *et al.* 2007). Thus, allelic variation at *Fls2* might be considered a QRL in a segregating population.

Fungal chitin also triggers basal resistance (Shibuya and Minami 2001). The host receptor-like kinase, *CERK1* (also called *LysM RLK1*), is a

component of chitin perception and defense signal transduction. Mutations in *CERK1* conditioned reduced resistance to the normally incompatible necrotrophic fungus *Alternaria brassicicola* (Miya *et al.* 2007), as well as quantitative differences in resistance to the biotrophic fungal pathogen *Erysiphe cichoracearum*, but not a bacterial pathogen *Pseudomonas syringea* (Ramonell *et al.* 2005; Wan *et al.* 2008). Thus, it seems probable that similar mutations or allelic changes might underlie QRL for fungal pathogens. These examples support the hypothesis that pattern recognition receptors acting in basal defense can condition quantitative differences in the resistance phenotype. The recognition of conserved pathogen features, such as flagellin or chitin, could also explain the broad spectrum resistance of some QRL (Appendix 1).

Hypothesis #3. QRLs are components of chemical warfare

Pathogen-produced phytotoxins have long been recognized as important compounds in promoting plant disease. The enzymes that detoxify them are recognized as important plant defenses, as are the "antibiotics" (phytoalexins) that plants deploy against pathogens. It is reasonable to expect that these types of compounds are components of QDR. Following QRL mapping in the *Arabidopsis-Botrytis* pathosystem (Denby *et al.* 2004), biochemical studies showed that levels of camalexin (a phytoalexin) were correlated with quantitative resistance, and that camalexin sensitivity of different pathogen isolates contributed to isolate specificity (Kliebenstein *et al.* 2005). Although this evidence is only correlative, it suggests that genes controlling phytoalexin levels underlie QRL.

The results of several studies using phytoalexin-deficient mutants (e.g., *pad2-1* and *pad3-1*; (Roetschi *et al.* 2001)) have suggested that reduced

amounts of glutathione was the cause of susceptibility to several pathogens (Parisy *et al.* 2007). Additional evidence connecting quantitative variation in host resistance to glutathione-mediated mechanisms was found in a bioinformatic analysis of rice, in which members of the glutathione *S*-transferase (GST) gene family were found to co-localize with QRL (Wisser *et al.* 2005). While GSTs have been implicated in diverse functions, their cytoprotective role during plant-pathogen interactions is well-documented (e.g., Refs. (Marrs 1996; Dean *et al.* 2005; Schlaeppi *et al.* 2008)).

Necrotrophic pathogens utilize numerous secondary metabolites to disrupt normal cell processes, trigger host cell death, and destroy host tissue. The well-studied pathogen *Botrytis cinerea* uses an arsenal of compounds to attack its hosts. Botrydial, a phytotoxin, has been shown to be a virulence factor (Colmenares *et al.* 2002). Likewise, production of oxalic acid is a potent weapon for *B. cinerea*, causing a reduction in host oxidative burst and defense responses, as well as producing an acidic environment for further enzymatic degradation (van Kan 2006). It has been shown that oxalate oxidase in the host can mitigate damage caused by pathogen-produced oxalic acid (Walz *et al.* 2008). Thus, the mode of necrotrophic pathogen attack naturally implicates toxin production and mitigation in quantitative disease resistance.

Hypothesis #4. QRL are involved in defense signal transduction

To modulate induced defense responses, plants have developed a complex system for the effective transmission of signals from initial pathogen perception to the activation of defense mechanisms. This often involves matching the defense response to the respective invader (e.g., biotrophic vs. necrotrophic pathogen vs. herbivore) and utilizes the phytohormones salicylic acid (SA), jasmonic acid (JA) and ethylene (ET). Different alleles of genes

involved in the regulation of these signaling pathways might be QRL. For example, mutant alleles of the transcription factor *WRKY33* increased susceptibility to the necrotrophic pathogen *B. cinerea* in *Arabidopsis* (Zheng *et al.* 2006). Similarly, mutations in *Arabidopsis* signaling component *MAP Kinase 4* resulted in heightened resistance to *Pseudomonas syringea* pv. *tomato* and *Peronospora parasitica* (Petersen *et al.* 2000). There are numerous other examples of signaling components that condition varying levels of increased susceptibility or resistance when mutated (e.g., Refs. (Kunkel and Brooks 2002; Koornneef and Pieterse 2008)). Often these signaling mutations increase resistance to a range of similar pathogens (i.e., biotrophic) but increase susceptibility to other pathogens (i.e., necrotrophs) and could also contribute to the broad spectrum resistance of some QRL (Appendix 1).

Hypothesis #5. QRLs are weak forms of R-genes

Several authors have posited that QRL are simply weaker forms of R-genes (Young 1996). There are several compelling lines of evidence that allelic variants at R-genes account for a proportion of QDR in plants. Co-localization of QRLs and R-genes has been noted in several species, including rice (Wang *et al.* 1994), maize (Xiao *et al.* 2007), and potato (Gebhardt and Valkonen 2001). In rice, both R-genes and R-gene analogues (RGAs; see glossary) were significantly associated with QRL (Wisser *et al.* 2005).

Although dogma has it that R-genes confer complete race-specific resistance and QRL confer partial race non-specific resistance, this is often not the case. There are numerous examples of R-genes that condition incomplete resistance (often defeated R-genes), including several that have been cloned and identified as NB-LRR genes. A catalytically impaired mutant

of *Xa21* confers partial rice blast resistance (Andaya and Ronald 2003). Other examples include genes from the maize common rust, flax rust, potato late blight, and tomato leaf mold systems (Lawrence *et al.* 1995; Parniske *et al.* 1997; Stewart *et al.* 2003; Smith and Hulbert 2005). Likewise, many QRL have been shown to be isolate- or race-specific (Appendix 2). *Rcg1*, a large-effect QRL for resistance to anthracnose stalk rot, has been isolated by map-based cloning and found to encode an NB-LRR resistance gene (Broglie *et al.* 2006). Although not known to be race specific, *Rcg1* illustrates the largely semantic distinction between R-genes and some QRL.

Thus, qualitative and quantitative disease resistance might only be two ends of a continuum with R-genes tending to lie toward one end of the spectrum and QRL toward the other (Figure 2). Although selection favors R-genes with strong effects, pathogen evolution can erode the effectiveness of R-genes, converting them into QRLs. It has been observed that when a pathogen strain overcomes an R-gene, the level of disease in the presence of the "defeated" R-gene (see glossary) is sometimes reduced relative to the level of disease in the absence of the R allele. This phenomenon, known as "residual resistance", is seen for the *Xa4* R-gene in the rice-*Xanthomonas oryzae* pv. *oryzae* system (Li *et al.* 1999) and also in the wheat stem rust (Brodny *et al.* 1986) and powdery mildew (Nass *et al.* 1981) pathosystems.

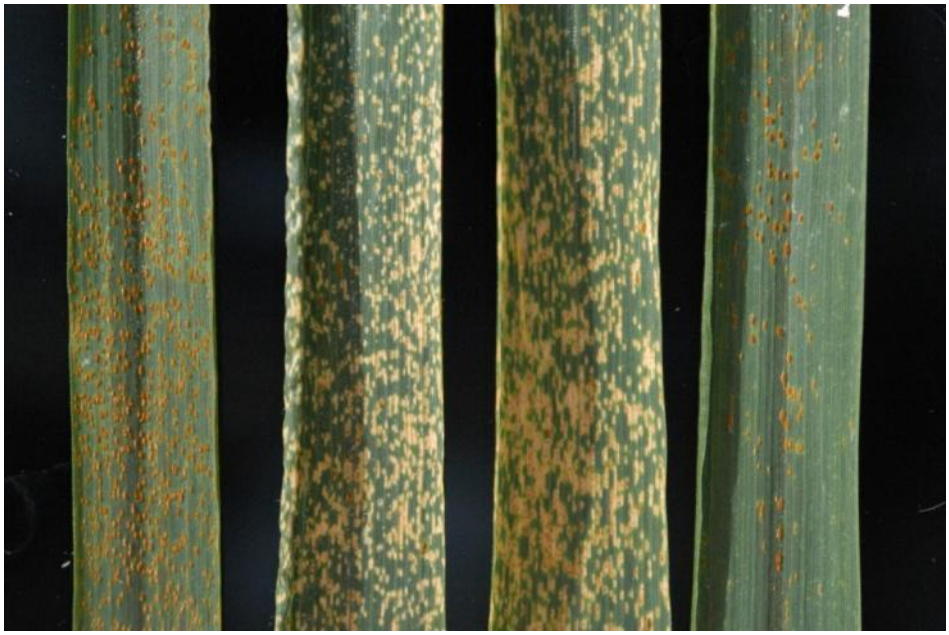


Figure 2. Different reaction types associated with R-genes in the wheat leaf rust pathosystem (causal agent, Puccinia triticina).

From left: Thatcher (susceptible) and Thatcher + *Lr12* (resistant), Thatcher + *Lr13* (resistant) and Thatcher + *Lr34* (resistant). The typical susceptible reaction of Thatcher is characterized by large number of uredinial pustules producing rust colored spores. The resistance genes *Lr12* and *Lr13* are characterized by a hypersensitive response (HR) that eliminates pathogen growth in an incompatible reaction. The resistance gene *Lr34* is not characterized by HR, but rather a reduction in disease severity. While *Lr34* can be distinguished as a single gene in a segregating population, lines carrying *Lr34* are compatible with *P. triticina* and do not show race specificity. *Lr34* tends to act as a large effect QRL blurring the distinction between R-genes and QRL. (*Photo reproduced with permission from James Kolmer, USDA-ARS*)

Hypothesis #6. QRLs are a unique set of previously un-identified genes

Three recently cloned QRL support the hypothesis that QRLs are a completely novel component of plant defense. In the wheat-leaf rust pathosystem, the resistance gene *Lr34* has been identified as a putitive adenosine triphosphate-binding cassette (ABC) transporter (Krattinger *et al.* 2009). A second gene in wheat, *Yr26*, was isolated by map-based cloning and shown to be a Kinase-START gene (Fu *et al.* 2009). The last QRL cloned to date is *pi21* from rice which confers resistance to rice blast. The recessive resistance of *pi21* results from the loss of a proline-rich gene (Fukuoka *et al.* 2009). A second QRL conditioning resistance to rice blast, *Pi34*, has been narrowed to a 65-kb region containing ten predicted open reading frames (Zenbayashi-Sawata *et al.* 2007). None of these candidate genes have sequence similarity to any previously reported defense genes.

With the exception of *Rcg1*, all of the QRLs that have been identified to date are unique genes that had not previously been associated with disease resistance. With the identification of additional QRL it is probable that new and novel classes of genes will be implicated in disease resistance. Research in QDR presents an exciting frontier in plant biology, where many unknowns wait to be uncovered. At the same time we must proceed with caution as previous knowledge may be of little use, as has aptly been demonstrated in the identification of these QRL. Particularly with interest of identifying candidate genes for further study or confirmation, previous research in plant disease resistance may have little bearing on the mechanisms of QDR.

The next frontier

Over the past several years, a detailed model of the gene-for-gene type of plant-pathogen interactions has emerged involving recognition, evasion and

defense (Nimchuk *et al.* 2003). Many facets of this model can be invoked as potential mechanisms of QDR, including variation in basal resistance, weak R-gene-mediated responses, differences in the speed and effectiveness of the defense response once the pathogen has been detected and even variable sensitivity to suppression of the defense responses by effectors. However, it appears probable that the molecular basis of QDR will draw upon an even broader mechanistic base. Aspects such as plant morphology and development, components of signal transduction systems, anti-microbial compounds such as phytoalexins and phytoanticipins, and other, as-yet unknown factors are also likely to be important components of QDR. Several lines of evidence also show that QDR is conditioned by genes previously unassociated with disease resistance that could control a range of morphological, detoxification, or developmental pathways in the plant host.

Two exciting frontiers in QDR research will be further isolation and characterization of QRL and the phenotypic analysis of QRLs as they affect the developmental biology and biochemistry of host-pathogen interactions. Cloning additional genes underlying QRL and determining their functions will reveal the ways in which QRLs contribute to plant defense. This knowledge will enable more efficient and effective utilization of these genes in crop improvement and protection. New genomic platforms in crop species, such as a nested association mapping population of maize, are providing unprecedented power for quantitative trait locus discovery and characterization (Yu *et al.* 2008; Buckler *et al.* 2009). With growing community resources, such as public recombinant inbred line populations (see glossary) and development of genome-wide association platforms, a greater focus on QDR in *Arabidopsis* could lead to additional advances in our understanding of this important agricultural and biological phenomenon.

The power of detailed observation will complement these genomic advances as careful phenotypic characterization will be essential to understanding these genes of modest effects. To distinguish minor phenotypic differences, analysis under natural conditions will be an important consideration (e.g., Ref. (Zipfel *et al.* 2004)). Valuable insights are being gained from observations on the resistance phenotypes (macroscopic and microscopic) associated with QRLs in mapping populations and near isogenic lines (see glossary). For example, characterization of QDR in the cereal rusts indicates that QRLs often act at the level of the cell wall, reducing the efficiency with which the biotrophic fungi enter the cell (Collins *et al.* 2007). An understanding of the developmental processes associated with pathogenesis, as well as the microscopic phenotypes associated with failure or attenuation of pathogenesis on resistant hosts, can provide important clues into the mechanisms of resistance. Current genomic advances situate QDR as an exciting field for systems biology research with the additional prospect of valuable application in crop improvement.

Glossary

Compatible interaction: a host-pathogen interaction that results in disease (the host is susceptible)

Defeated R-gene: a resistance gene that has become ineffective

Genetic architecture: the number, effect and genomic distribution of loci affecting a given trait

Incompatible interaction: a host-pathogen interaction that does not result in disease (the host is resistant)

NB-LRR (Nucleotide-binding leucine rich repeat): two amino acid sequence motifs commonly found in resistance genes

Near isogenic lines (NIL): two inbred lines that differ at only a small genomic region

Pathosystem: the combination of a specific host species and pathogen species

Pattern recognition receptors: proteins that identify molecules associated with microbial pathogens such as flagellin or chitin components

Quantitative disease resistance (QDR): resistance that is expressed as a reduction in disease, rather than as the absence of disease

Quantitative trait locus (**QTL)**: a locus with an effect on a quantitative trait (i.e., a trait showing continuous variation)

Quantitative resistance locus (QRL): a locus with an effect on QDR

R-gene breakdown: the phenomenon of a resistance gene becoming ineffective in a crop variety

Resistance breakdown: the phenomenon of a resistant cultivar becoming susceptible due to changes in the pathogen race

Resistance gene analogs: putative genes that share sequence similarity with known R- genes

Recombinant Inbred Line (RIL): an inbred line produced from an initial cross followed by continuous inbreeding; populations of RILs are often used for QTL mapping studies

APPENDIX 1.

Broad-spectrum QRLs: evidence for multiple disease resistance

In both natural and agricultural systems, plants are exposed to taxomically diverse pathogens, including fungi, bacteria, oomycetes, viruses, viroids and nematodes. For agricultural production, plant cultivars should therefore exhibit acceptable levels of resistance to a spectrum of pathogens, in order to prevent economic yield losses. Indeed, crop varieties showing multiple disease resistance have been developed through phenotypic selection. This raises a fundamental question: do single loci with pleiotropic effects on multiple diseases contribute to that spectrum of resistance?

Various lines of evidence support the existence of multiple disease resistance genes in plants. Several studies, mostly conducted using induced mutants and transformants of *Arabidopsis*, have implicated specific genes that condition multiple disease resistance (e.g., *NDR1* (Century *et al.* 1997), *NPR1* (Mou *et al.* 2003), *cir1, 2,& 3* (Murray *et al.* 2005), *esa1* (Tierens *et al.* 2002), *WRKY33* (Zheng *et al.* 2006), and *AS1* (Nurmberg *et al.* 2007)). In several of these studies, multiple disease resistance genes were reported to be effective against one group of pathogens (e.g., biotrophs), while increasing the susceptibility of the plant to another group (e.g., necrotrophs). Currently, it is not known whether these genes contribute to natural variation in resistance that could be manipulated by selection.

Further evidence for multiple disease resistance genes based on quantitative genetic analysis comes in several forms. Phenotypic and genetic correlations for resistance to different pathogens have been documented in various populations (e.g., Refs. (Hill and Leath 1975; Mitchell-Olds *et al.* 1995)). In populations under selection for QDR, correlated reductions in diseases for which resistance was not specifically selected have also been

reported (Lambert and White 1997). From QRL mapping studies for disease resistance, QRL for different diseases often co-localize and clusters of QRL for multiple diseases have been observed in summaries of QRL mapping studies (Wisser *et al.* 2005; Wisser *et al.* 2006). In some cases, alleles for resistance at co-localizing rice QRL for different diseases were inherited from the same parent (Wisser *et al.* 2005). The resolution of these studies, however, does not permit pleiotropy to be distinguished from linkage, limiting inference about quantitative variation in multiple disease resistance genes. Using a collection of diverse maize inbred lines with gene-level linkage disequilibrium (a determinant of mapping resolution) (Flint-Garcia *et al.* 2005), significant positive genetic correlations were found for resistance to three different foliar fungal pathogens (R.Wisser *et al.*, unpubished), suggesting the existence of functionally variable multiple disease resistance genes.

APPENDIX 2.

Narrow-spectrum QRLs: evidence for gene-for-gene interactions in quantitative disease resistance

Although QDR is generally presumed to be race non-specific (e.g., (van der Plank 1968)), several lines of evidence challenge this dogma. Race-specific QRL have been identified in multiple pathosystems. These include rose blackrot (*Diplocarpon rosae*) (Whitaker *et al.* 2007); rice blast (*Magnaporthe oryzae*) (Li *et al.* 2006); leaf rust in barley (*Puccinia hordei*) (Parlevliet 1978); vascular wilt of melon (*Fusarium oxysporum*) (Perchepied *et al.* 2005); black stem in sunflower (*Phoma macdonaldii*) (Darvishzadeh *et al.* 2007); and leaf stripe in barley (*Pyrenophora graminea*) (Arru *et al.* 2003). In the rice-*Xanthomonas* pathosystem, for example, Li *et al.* (Li *et al.* 2006) evaluated two mapping populations for resistance to ten different pathogen races and found that numerous QRL were effective only against a sub-set of the pathogen races. Even more striking was that, for some of the QRL, resistance was contributed by one parent for a given race and the other parent for a different race. This study provides clear examples of the race-specificity of QRL and the authors speculate that QDR is a weaker form of race-specific (R-gene mediated) resistance (Li *et al.* 2006).

Pathogen adaption on hosts with partial resistance has also been demonstrated in several pathosystems. A population of *Cochliobolus heterostrophus*, the causal agent of southern leaf blight in maize, was shown to have increased virulence on the specific host genotype on which increased virulence was selected compared with host genotypes on which the pathogen population was not selected (Kolmer and Leonard 1986). Likewise, isolates of *Phytophthora infestans*, the causal agent of potato late blight, were shown to be more aggressive on the cultivar from which they were isolated (Bjor and

Mulelid 1991). Both of these studies considered cultivars with only QDR, supporting the hypothesis that QRL condition resistance in a race-specific manner.

This evidence is consistent with the hypothesis that some QRL condition a weaker form of R-gene mediated defense, as proposed by Parlevliet and Zadoks (Parlevliet and Zadoks 1977), though it is also possible that other mechanisms provide isolate-specific resistance (e.g., camalexin sensitivity). In this view, phenotypic variance and durability can be explained by a minor gene-for-minor gene interaction where virulence genes of minor effect in the pathogen correspond to resistance genes of minor effect in the host (QRL). This is supported by cultivar x isolate interactions and/or race specificity of QRL (Marcel *et al.* 2008). A rice blast QRL, *Pi34*, and a corresponding avirulence (aggressiveness) gene, *AVR-Pi34*, have been shown to interact in a typical gene-for-gene manner, giving further credibility to the idea of this type of minor interaction (Zenbayashi-Sawata *et al.* 2005).

REFERENCES

Albar, L., M. Lorieux, N. Ahmadi, I. Rimbault, A. Pinel, A. A. Sy, D. Fargette and A. Ghesquière (1998). Genetic Basis and Mapping of the Resistance to Rice Yellow Mottle Virus. I. QTLs Identification and Relationship between Resistance and Plant Morphology. *Theoretical and Applied Genetics* **97** (7): 1145-1154.

Andaya, C. B. and P. C. Ronald (2003). A Catalytically Impaired Mutant of the Rice *Xa21* Receptor Kinase Confers Partial Resistance to *Xanthomonas oryzae* Pv *Oryzae*. *Physiological and Molecular Plant Pathology* **62** (4): 203-208.

Arru, L., E. Francia and N. Pecchioni (2003). Isolate-Specific QTLs of Resistance to Leaf Stripe (*Pyrenophora Graminea*) in the 'Steptoe' x 'Morex' Spring Barley Cross. *Theoretical and Applied Genetics* **106** ((4):): 668-675.

Ayliffe, M., R. Singh and E. Lagudah (2008). Durable Resistance to Wheat Stem Rust Needed. *Current Opinion in Plant Biology* **11** (2): 187-192.

Balint-Kurti, P. J., J. C. Zwonitzer, M. E. Pe, G. Pea, M. Lee and A. J. Cardinal (2008). Identification of Quantitative Trait Loci for Resistance to Southern Leaf Blight and Days to Anthesis in Two Maize Recombinant Inbred Line Populations. *Phytopathology* **98** (3): 315-320.

Bent, A. F. and D. Mackey (2007). Elicitors, Effectors, and R Genes: The New Paradigm and a Lifetime Supply of Questions. *Annual Review of Phytopathology* **45** (1): 399-436.

Bjor, T. and K. Mulelid (1991). Differential Resistance to Tuber Late Blight in Potato Cultivars without R-Genes. *Potato Research* **34** (1): 3-8.

Bradley, D. J., G. S. Gilbert and I. M. Parker (2003). Susceptibility of Clover

Species to Fungal Infection: The Interaction of Leaf Surface Traits and Environment. *Am. J. Bot.* **90** (6): 857-864.

Brandwagt, B. F., T. J. A. Kneppers, H. J. J. Nijkamp and J. Hille (2002). Overexpression of the Tomato Asc-1 Gene Mediates High Insensitivity to Aal Toxins and Fumonisin B1 in Tomato Hairy Roots and Confers Resistance to *Alternaria alternata* F. Sp. *Lycopersici* in *Nicotiana Umbratica* Plants. *Molecular Plant-Microbe Interactions* **15** (1): 35-42.

Brodny, U., R. R. Nelson and G. L.V. (1986). Residual and Interactive Expressions of" Defeated" Wheat Stem Rust Resistance Genes. *Phytopathology* **76** (5): 546-549.

Broglie, K. E., K. H. Butler, C. A. De Silva, T. J. Frey, J. A. Hawk, D. S. Multani, C. Wolters and J. C. Petra (2006). Polynucleotides and Methods for Making Plants Resistant to Fungal Pathogens. United States, E.I. du Pont de Nemours and Company, Pioneer Hi-Bred International, Inc.; University of Delaware. **United States Patent 20060223102**.

Bryan, W. P., S. J. Gurmukh and D. M. Michael (2004). A Major Suppressor of Cell Death, *SLM1*, Modifies the Expression of the Maize (Zea Mays L.) Lesion Mimic Mutation *Les23*. *Genome* **47**: 961-969.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J. C. Glaubitz, M. M. Goodman, C. Harjes, K. Guill, D. E. Kroon, S. Larsson, N. K. Lepak, H. Li, S. E. Mitchell, G. Pressoir, J. A. Peiffer, M. O. Rosas, T. R. Rocheford, M. C. Romay, S. Romero, S. Salvo, H. S. Villeda, H. Sofia da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich and M. D. McMullen (2009). The Genetic Architecture of Maize Flowering Time. *Science* **325** (5941): 714-718.

Büschges, R., K. Hollricher, R. Panstruga, G. Simons, M. Wolter, A. Frijters, R. van Daelen, T. van der Lee, P. Diergaarde, J. Groenendijk, S. Töpsch, P. Vos, F. Salamini and P. Schulze-Lefert (1997). The Barley *Mlo* Gene: A Novel Control Element of Plant Pathogen Resistance. *Cell* **88** (5): 695-705.

Century, K. S., A. D. Shapiro, P. P. Repetti, D. Dahlbeck, E. Holub and B. J. Staskawicz (1997). *NDR1*, a Pathogen-Induced Component Required for *Arabidopsis* Disease Resistance. *Science* **278** (5345): 1963-1965.

Collins, A., D. Milbourne, L. Ramsay, R. Meyer, C. Chatot-Balandras, P. Oberhagemann, W. De Jong, C. Gebhardt, E. Bonnel and R. Waugh (1999). QTL for Field Resistance to Late Blight in Potato Are Strongly Correlated with Maturity and Vigour. *Molecular Breeding* **5** (5): 387-398.

Collins, N. C., R. E. Niks and P. Schulze-Lefert (2007). Resistance to Cereal Rusts at the Plant Cell Wall - What Can We Learn from Other Host-Pathogen Systems? *Australian Journal of Agricultural Research* **58** (6): 476-489.

Colmenares, A. J., J. Aleu, R. Durán-Patrón, I. G. Collado and R. Hernández-Galán (2002). The Putative Role of Botrydial and Related Metabolites in the Infection Mechanism of *Botrytis Cinerea*. *Journal of Chemical Ecology* **28** (5): 997-1005.

Darvishzadeh, R., S. P. Kiani, G. Dechamp-Guillaume, L. Gentzbittel and A. Sarrafi (2007). Quantitative Trait Loci Associated with Isolate Specific and Isolate Nonspecific Partial Resistance to *Phoma Macdonaldii* in Sunflower. *Plant Pathology* **56** (5): 855-861.

Dean, J. D., P. H. Goodwin and T. Hsiang (2005). Induction of Glutathione S-Transferase Genes of *Nicotiana Benthamiana* Following Infection by *Colletotrichum Destructivum* and *C. Orbiculare* and Involvement of One

in Resistance. *J. Exp. Bot.* **56** (416): 1525-1533.

Denby, K. J., P. Kumar and D. J. Kliebenstein (2004). Identification of *Botrytis Cinerea* Susceptibility Loci in *Arabidopsis Thaliana*. *Plant Journal* **38** (3): 473-486.

Dunning, F. M., W. Sun, K. L. Jansen, L. Helft and A. F. Bent (2007). Identification and Mutational Analysis of Arabidopsis *Fls2* Leucine-Rich Repeat Domain Residues That Contribute to Flagellin Perception. *Plant Cell*: tpc.106.048801.

Flint-Garcia, S. A., A.-C. Thuillet, J. Yu, G. Pressoir, S. M. Romero, S. E. Mitchell, J. Doebley, S. Kresovich, M. M. Goodman and E. S. Buckler (2005). Maize Association Population: A High-Resolution Platform for Quantitative Trait Locus Dissection. *Plant Journal* **44** (6): 1054-1064.

Fu, D., C. Uauy, A. Distelfeld, A. Blechl, L. Epstein, X. Chen, H. Sela, T. Fahima and J. Dubcovsky (2009). A Kinase-Start Gene Confers Temperature-Dependent Resistance to Wheat Stripe Rust. *Science* **323** (5919): 1357-1360.

Fukuoka, S., N. Saka, H. Koga, K. Ono, T. Shimizu, K. Ebana, N. Hayashi, A. Takahashi, H. Hirochika, K. Okuno and M. Yano (2009). Loss of Function of a Proline-Containing Protein Confers Durable Disease Resistance in Rice. *Science* **325** (5943): 998-1001.

Gebhardt, C. and J. P. T. Valkonen (2001). Organization of Genes Controlling Disease Resistance in the Potato Genome. *Annual Review of Phytopathology* **39** (1): 79-102.

Hill, R. R. and K. T. Leath (1975). Genotypic and Phenotypic Correlations for Reaction to Five Foliar Pathogens in Alfalfa. *Theoretical and Applied Genetics* **45** (6): 254-258.

Hu, G., C. A. Webb and S. H. Hulbert (1997). Adult Plant Phenotype of the

Rp1-Dj Compound Rust Resistance Gene in Maize. *Phytopathology* **87** (3): 236.

Johal, G. S. and S. P. Briggs (1992). Reductase Activity Encoded by the *Hm1* Disease Resistance Gene in Maize. *Science* **258** (5084): 985-987.

Jones, J. D. G. and J. L. Dangl (2006). The Plant Immune System. *Nature* **444** (7117): 323-329.

Kamoun, S., E. Huitema and V. G. A. A. Vleeshouwers (1999). Resistance to *Oomycetes*: A General Role for the Hypersensitive Response? *Trends in Plant Science* **4** (5): 196-200.

Kliebenstein, D. J., H. C. Rowe and K. J. Denby (2005). Secondary Metabolites Influence *Arabidopsis/Botrytis* Interactions: Variation in Host Production and Pathogen Sensitivity. *Plant Journal* **44** (1): 25-36.

Kolmer, J. A. and K. L. Leonard (1986). Genetic Selection and Adaptation of *Cochliobolus Heterostrophus* to Corn Hosts with Partial Resistance. *Phytopathology* **76** (8): 774-777.

Koornneef, A. and C. M. J. Pieterse (2008). Cross Talk in Defense Signaling. *Plant Physiology* **146** (3): 839-844.

Kover, P. X. and B. A. Schaal (2002). Genetic Variation for Disease Resistance and Tolerance among *Arabidopsis Thaliana* Accessions. *Proceedings of the National Academy of Sciences, USA* **99** (17): 11270-11274.

Krattinger, S. G., E. S. Lagudah, W. Spielmeyer, R. P. Singh, J. Huerta-Espino, H. McFadden, E. Bossolini, L. L. Selter and B. Keller (2009). A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. *Science* **323** (5919): 1360-1363.

Kunkel, B. N. and D. M. Brooks (2002). Cross Talk between Signaling Pathways in Pathogen Defense. *Current Opinion in Plant Biology* **5** (4):

325-331.

Lambert, R. J. and D. G. White (1997). Disease Reaction Changes from Tandem Selection for Multiple Disease Resistance in Two Maize Synthetics. *Crop Science* **37** (1): 66-69.

Landeo, J., M. Gastelo, H. Pinedo and F. Flores (1995). Breeding for Horizontal Resistance to Late Blight in Potato Free of R-Genes. *Phytophthora infestans*. L. Dowley, E. Bannon, L. Cooke, T. Keane and E. O'Sullivan. Dublin, Ireland, Bole Press Ltd**:** 268-274.

Lawrence, G. J., E. J. Finnegan, M. A. Ayliffe and J. G. Ellis (1995). The L6 Gene for Flax Rust Resistance Is Related to the *Arabidopsis* Bacterial Resistance Gene *RPS2* and the Tobacco Viral Resistance Gene N. *The Plant Cell* **7** (8): 1195-1206.

Li, Z. K., M. Arif, D. B. Zhong, B. Y. Fu, J. L. Xu, J. Domingo-Rey, J. Ali, C. H. M. Vijayakumar, S. B. Yu and G. S. Khush (2006). Complex Genetic Networks Underlying the Defensive System of Rice (*Oryza Sativa* L.) to *Xanthomonas oryzae* Pv. *Oryzae*. *Proceedings of the National Academy of Sciences, USA* **103** (21): 7994-7999.

Li, Z. K., L. J. Luo, H. W. Mei, A. H. Paterson, X. H. Zhao, D. B. Zhong, Y. P. Wang, X. Q. Yu, L. Zhu, R. Tabien, J. W. Stansel and C. S. Ying (1999). A "Defeated" Rice Resistance Gene Acts as a QTL against a Virulent Strain of *Xanthomonas Oryzae* Pv. *Oryzae*. *Molecular and General Genetics* **261** (1): 58-63.

Liu, B., S. Zhang, X. Zhu, Q. Yang, S. Wu, M. Mei, R. Mauleon, J. Leach, T. Mew and H. Leung (2004). Candidate Defense Genes as Predictors of Quantitative Blast Resistance in Rice. *Molecular Plant-Microbe Interactions* **17** (10): 1146-1152.

Lorang, J. M., T. A. Sweat and T. J. Wolpert (2007). Plant Disease

Susceptibility Conferred by A "Resistance" Gene. *Proceedings of the National Academy of Sciences, USA* **104** (37): 14861-14866.

Marcel, T. C., B. Gorguet, M. Truong Ta, Z. Kohutova, A. Vels and R. E. Niks (2008). Isolate Specificity of Quantitative Trait Loci for Partial Resistance of Barley to *Puccinia Hordei* Confirmed in Mapping Populations and near-Isogenic Lines. *New Phytologist* **177** (3): 743-755.

Marrs, K. A. (1996). The Functions and Regulation of Glutathione S-Transferases in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47** (1): 127-158.

McDonald, B. A. and C. Linde (2002). Pathogen Population Genetics, Evolutionary Potential, and Durable Resistance. *Annual Review of Phytopathology* **40** (1): 349-379.

Melotto, M., W. Underwood, J. Koczan, K. Nomura and S. Y. He (2006). Plant Stomata Function in Innate Immunity against Bacterial Invasion. *Cell* **126** (5): 969-980.

Mitchell-Olds, T., R. V. James, M. J. Palmer and P. H. Williams (1995). Genetics of *Brassica Rapa* (Syn. *Campestris*). 2. Multiple Disease Resistance to Three Fungal Pathogens: *Peronospora Parasitica, Albugo Candida* and *Leptosphaeria Maculans. Heredity* **75** (4): 362-369.

Miya, A., P. Albert, T. Shinya, Y. Desaki, K. Ichimura, K. Shirasu, Y. Narusaka, N. Kawakami, H. Kaku and N. Shibuya (2007). Cerk1, a Lysm Receptor Kinase, is Essential for Chitin Elicitor Signaling in *Arabidopsis*. *Proceedings of the National Academy of Sciences, USA* **104** (49): 19613-19618.

Mou, Z., W. Fan and X. Dong (2003). Inducers of Plant Systemic Acquired

Resistance Regulate *NPR1* Function through Redox Changes. *Cell* **113** (7): 935-944.

Murray, S. L., N. Adams, D. J. Kliebenstein, G. J. Loake and K. J. Denby (2005). A Constitutive *PR-1*::Luciferase Expression Screen Identifies *Arabidopsis* Mutants with Differential Disease Resistance to Both Biotrophic and Necrotrophic Pathogens. *Molecular Plant Pathology* **6** (1): 31-41.

Nagy, E., T.-C. Lee, W. Ramakrishna, Z. Xu, P. Klein, P. SanMiguel, C.-P. Cheng, J. Li, K. Devos, K. Schertz, L. Dunkle and J. Bennetzen (2007). Fine Mapping of the Pc Locus of Sorghum Bicolor , a Gene Controlling the Reaction to a Fungal Pathogen and Its Host-Selective Toxin. *Theoretical and Applied Genetics* **114** (6): 961-970.

Nass, H. A., W. L. Pedersen, D. R. Mackenzie and R. R. Nelson (1981). The Residual Effects of Some Defeated Powdery Mildew *Erysiphe-Graminis-F-Sp-Tritici* Resistance Genes in Isolines of Winter Wheat. *Phytopathology* **71** (12): 1315-1318.

Nimchuk, Z., T. Eulgem, B. F. Holt Iii and J. L. Dangl (2003). Recognition and Response in the Plant Immune System. *Annual Review of Genetics* **37** (1): 579-609.

Nurmberg, P. L., K. A. Knox, B.-W. Yun, P. C. Morris, R. Shafiei, A. Hudson and G. J. Loake (2007). The Developmental Selector *AS1* Is an Evolutionarily Conserved Regulator of the Plant Immune Response. *Proceedings of the National Academy of Sciences, USA* **104** (47): 18795-18800.

Parisy, V., B. Poinssot, L. Owsianowski, A. Buchala, J. Glazebrook and F. Mauch (2007). Identification of *PAD2* as a G-Glutamylcysteine Synthetase Highlights the Importance of Glutathione in Disease

Resistance of Arabidopsis. *The Plant Journal* **49** (1): 159-172.

Parlevliet, J. (2002). Durability of Resistance against Fungal, Bacterial and Viral Pathogens; Present Situation. *Euphytica* **124** (2): 147-156.

Parlevliet, J. E. (1978). Race-Specific Aspects of Polygenic Resistance of Barley to Leaf Rust, *Puccinia Hordei*. *European Journal of Plant Pathology* **84** (4): 121-126.

Parlevliet, J. E. and J. C. Zadoks (1977). The Integrated Concept of Disease Resistance: A New View Including Horizontal and Vertical Resistance in Plants. *Euphytica* **26** (1): 5-21.

Parniske, M., K. E. Hammond-Kosack, C. Golstein, C. M. Thomas, D. A. Jones, K. Harrison, B. B. H. Wulff and J. D. G. Jones (1997). Novel Disease Resistance Specificities Result from Sequence Exchange Between Tandemly Repeated Genes at the Cf-4/9 Locus of Tomato. *Cell* **91** (6): 821-832.

Perchepied, L., C. Dogimont and M. Pitrat (2005). Strain-Specific and Recessive QTLs Involved in the Control of Partial Resistance to *Fusarium oxysporum* F. Sp. *Melonis* Race 1.2 in a Recombinant Inbred Line Population of Melon. *Theoretical and Applied Genetics* **111** (1): 65-74.

Petersen, M., P. Brodersen, H. Naested, E. Andreasson, U. Lindhart, B. Johansen, H. B. Nielsen, M. Lacy, M. J. Austin, J. E. Parker, S. B. Sharma, D. F. Klessig, R. Martienssen, O. Mattsson, A. B. Jensen and J. Mundy (2000). *Arabidopsis* Map Kinase 4 Negatively Regulates Systemic Acquired Resistance. *Cell* **103** (7): 1111-1120.

Ramonell, K., M. Berrocal-Lobo, S. Koh, J. Wan, H. Edwards, G. Stacey and S. Somerville (2005). Loss-of-Function Mutations in Chitin Responsive Genes Show Increased Susceptibility to the Powdery Mildew Pathogen

Erysiphe Cichoracearum. *PLANT PHYSIOLOGY* **138** (2): 1027-1036.

Roetschi, A., A. Si-Ammour, L. Belbahri, F. Mauch and B. Mauch-Mani (2001). Characterization of an *Arabidopsis-Phytophthora* Pathosystem: Resistance Requires a Functional *PAD2* Gene and Is Independent of Salicylic Acid, Ethylene and Jasmonic Acid Signalling. *The Plant Journal* **28** (3): 293-305.

Romer, P., S. Hahn, T. Jordan, T. Strauss, U. Bonas and T. Lahaye (2007). Plant Pathogen Recognition Mediated by Promoter Activation of the Pepper *Bs3* Resistance Gene. *Science* **318** (5850): 645-648.

Schlaeppi, K., N. Bodenhausen, A. Buchala, F. Mauch and P. Reymond (2008). The Glutathione-Deficient Mutant *PAD2-1* Accumulates Lower Amounts of Glucosinolates and Is More Susceptible to the Insect Herbivore *Spodoptera Littoralis*. *The Plant Journal* **55** (5): 774-786.

Shibuya, N. and E. Minami (2001). Oligosaccharide Signaling for Defense Responses in Plants. *Physiological and Molecular Plant Pathology* **59** (5): 223-233.

Sindhu, A., S. Chintamanani, A. S. Brandt, M. Zanis, S. R. Scofield and G. S. Johal (2008). A Guardian of Grasses: Specific Origin and Conservation of a Unique Disease-Resistance Gene in the Grass Lineage. *Proceedings of the National Academy of Sciences, USA*: 0711406105.

Singh, R. P., J. Huerta-Espino and H. M. William (2005). Genetics and Breeding for Durable Resistance to Leaf and Stripe Rusts in Wheat. *Turkish Journal of Agriculture Forestry* **29**: 121-127.

Smith, S. M. and S. H. Hulbert (2005). Recombination Events Generating a Novel *Rp1* Race Specificity. *Molecular Plant-Microbe Interactions* **18** (3): 220-228.

Steffenson, B. J., P. M. Hayes and A. Kleinhofs (1996). Genetics of Seedling

and Adult Plant Resistance to Net Blotch (*Pyrenophora Teres F. Teres*) and Spot Blotch (*Cochliobolus Sativus*) in Barley. *Theoretical and Applied Genetics* **92** (5): 552-558.

Stewart, H. E., J. E. Bradshaw and B. Pande (2003). The Effect of the Presence of R-Genes for Resistance to Late Blight (*Phytophthora infestans*) of Potato (*Solanum tuberosum*) on the Underlying Level of Field Resistance. *Plant Pathology* **52** (2): 193-198.

Stokstad, E. (2007). Plant Pathology: Deadly Wheat Fungus Threatens World's Breadbaskets. *Science* **315** (5820): 1786-1787.

Thompson, D. L. and R. R. Bergquist (1984). Inheritance of Mature Plant Resistance to *Helminthosporium maydis* Race 0 in Maize *Crop Sci* **24** (4): 807-811.

Tierens, K. F. M. J., B. P. H. J. Thomma, R. P. Bari, M. Garmier, K. Eggermont, M. Brouwer, I. A. M. A. Penninckx, W. F. Broekaert and B. P. A. Cammue (2002). *Esa1*, an *Arabidopsis* Mutant with Enhanced Susceptibility to a Range of Necrotrophic Fungal Pathogens, Shows a Distorted Induction of Defense Responses by Reactive Oxygen Generating Compounds. *Plant Journal* **29** (2): 131-140.

van der Plank, J. E. (1968). Disease Resistance in Plants. London, Academic Press.

van Kan, J. A. L. (2006). Licensed to Kill: The Lifestyle of a Necrotrophic Plant Pathogen. *Trends in Plant Science* **11** (5): 247-253.

Walz, A., I. Zingen-Sell, M. Loeffler and M. Sauer (2008). Expression of an Oxalate Oxidase Gene in Tomato and Severity of Disease Caused by *Botrytis Cinerea* and *Sclerotinia Sclerotiorum*. *Plant Pathology* **57** (3): 453-458.

Wan, J., X.-C. Zhang, D. Neece, K. M. Ramonell, S. Clough, S.-y. Kim, M. G.

Stacey and G. Stacey (2008). A *LysM* Receptor-Like Kinase Plays a Critical Role in Chitin Signaling and Fungal Resistance in *Arabidopsis*. *The Plant Cell* **20** (2): 471-481.

Wang, G. L., D. J. Mackill, J. M. Bonman, S. R. McCouch, M. C. Champoux and R. J. Nelson (1994). RFLP Mapping of Genes Conferring Complete and Partial Resistance to Blast in a Durably Resistant Rice Cultivar. *Genetics* **136** (4): 1421-1434.

Wang, Z., G. Taramino, D. Yang, G. Liu, S. V. Tingey, G. H. Miao and G. L. Wang (2001). Rice ESTs with Disease-Resistance Gene- or Defense-Response Gene-Like Sequences Mapped to Regions Containing Major Resistance Genes or QTLs. *Molecular Genetics and Genomics* **265** (2): 302-310.

Whitaker, V. M., K. Zuzek and S. C. Hokanson (2007). Resistance of 12 Rose Genotypes to 14 Isolates of *Diplocarpon Rosae* Wolf (Rose Blackspot) Collected from Eastern North America. *Plant Breeding* **126** (1): 83-88.

Wilson, I. W., C. L. Schiff, D. E. Hughes and S. C. Somerville (2001). Quantitative Trait Loci Analysis of Powdery Mildew Disease Resistance in the *Arabidopsis Thaliana* Accession Kashmir-1. *Genetics* **158** (3): 1301-1309.

Wisser, R. J., P. J. Balint-Kurti and R. J. Nelson (2006). The Genetic Architecture of Disease Resistance in Maize: A Synthesis of Published Studies. *Phytopathology* **96** (2): 120-129.

Wisser, R. J., Q. Sun, S. H. Hulbert, S. Kresovich and R. J. Nelson (2005). Identification and Characterization of Regions of the Rice Genome Associated with Broad-Spectrum, Quantitative Disease Resistance. *Genetics* **169** (4): 2277-2293.

Xiao, W., J. Zhao, S. Fan, L. Li, J. Dai and M. Xu (2007). Mapping of Genome-

Wide Resistance Gene Analogs (RGAs) in Maize (*Zea mays* L.). *Theoretical and Applied Genetics* **115** (4): 501-508.

Young, N. D. (1996). QTL Mapping and Quantitative Disease Resistance in Plants. *Annual Review of Phytopathology* **34** (1): 479-501.

Yu, J., J. B. Holland, M. D. McMullen and E. S. Buckler (2008). Genetic Design and Statistical Power of Nested Association Mapping in Maize. *Genetics* **178** (1): 539-551.

Zenbayashi-Sawata, K., T. Ashizawa and S. Koizumi (2005). *Pi34-AvrPi34*: A New Gene-for-Gene Interaction for Partial Resistance in Rice to Blast Caused by *Magnaporthe Grisea*. *Journal of General Plant Pathology* **71** (6): 395-401.

Zenbayashi-Sawata, K., S. Fukuoka, S. Katagiri, M. Fujisawa, T. Matsumoto, T. Ashizawa and S. Koizumi (2007). Genetic and Physical Mapping of the Partial Resistance Gene, *Pi34*, to Blast in Rice. *Phytopathology* **97** (5): 598-602.

Zheng, Z., S. A. Qamar, Z. Chen and T. Mengiste (2006). Arabidopsis *WRKY33* Transcription Factor Is Required for Resistance to Necrotrophic Fungal Pathogens. *The Plant Journal* **48** (4): 592-605.

Zhu, H., L. Gilchrist, P. Hayes, A. Kleinhofs, D. Kudrna, Z. Liu, L. Prom, B. Steffenson, T. Toojinda and H. Vivar (1999). Does Function Follow Form? Principal Qtls for Fusarium Head Blight (FHB) Resistance Are Coincident with QTLs for Inflorescence Traits and Plant Height in a Doubled-Haploid Population of Barley. *Theoretical and Applied Genetics* **99** (7): 1221-1232.

Zipfel, C., S. Robatzek, L. Navarro, E. J. Oakeley, J. D. G. Jones, G. Felix and T. Boller (2004). Bacterial Disease Resistance in *Arabidopsis* through Flagellin Perception. *Nature* **428** (6984): 764-767.

CHAPTER 2

Uncovering the genetic basis for complex disease resistance in maize using nested association mapping

ABSTRACT

Understanding the genetics of resistance to plant pathogens has implications for food production, food security, and food safety as well as importance in the fields of plant biology and evolutionary biology. Research in plant disease resistance has led to the identification of numerous resistance genes, most of which condition complete resistance in the host when challenged with an avirulent pathogen. Recent research has begun to identify the genes of small effects responsible for complex disease resistance in plants. Here we describe the genetic architecture of complex resistance in maize for an economically important disease, northern leaf blight (NLB) and implicate several genes through genome-wide nested association mapping. We evaluated a maize reference design population, consisting of 4,630 recombinant inbred lines, over three environments for quantitative resistance to NLB, giving highly heritable resistance phenotypes. We identified over 220 resistance alleles at 30 different quantitative trait loci (QTL). Most QTL identified had a range of estimated allele effects. We show that large variation in the resistance phenotype is the result of the accumulation of numerous loci of small additive effects. This genetic architecture is very similar to that of flowering time in maize, a classic complex trait. Genome-wide nested association mapping identified genes at six of the QTL that have been associated with disease resistance. Three QTL were associated with single

nucleotide polymorphisms (SNPs) at genes containing RLK–LRR domains, two QTL were associated with SNPs at ethylene response factors, and one QTL was associated with an *Mlo*-like gene. This report links these defense-related genes with complex disease resistance in plants. Based on this work and concurrent research, we consider maize an excellent model system for the study of complex plant disease resistance.

INTRODUCTION

Complex disease resistance in plants (commonly referred to as quantitative disease resistance; QDR) has importance in development of durably resistant plant cultivars (Ayliffe *et al.* 2008). The complete resistance conferred by single dominant resistance genes tends to be less durable than QDR, being subject to "break-down" due to evolution of pathogen populations. Resistance genes of large effect are typically involved in direct or indirect recognition of the pathogen. In the pathogen, mutation or loss of the recognition target renders the resistance genes ineffective. QDR, on the other hand, is conditioned by numerous genes of small partial effect, leading to more durable resistance, as loss of a single gene does not leave the host completely susceptible.

Race-specific resistance genes have been found to largely fall into few classes of genes, the most common being nucleotide binding – leucine rich repeat (NB-LRR). Due to the unique nature of QDR, however, it is expected that new classes of genes previously unassociated with plant disease resistance will contribute to this important phenotype. Indeed, three genes that condition partial race-non-specific resistance have recently been identified and found to encode novel types of defense related genes. A gene conferring temperature dependent partial resistance to stripe rust in wheat, *Yr36*, was cloned and found to be a kinase with a putative START lipid-binding domain (Fu *et al.* 2009). A second gene conferring resistance to wheat leaf rust, *Lr34*, was cloned and found to be a putative ABC transporter (Krattinger *et al.* 2009). *Lr34* also confers resistance to other fungal diseases, stripe rust and powdery mildew, indicating a role in general defense against biotrophic pathogens. The loss of a proline-rich protein that includes a putative heavy metal–binding

domain was shown to underlie the disease resistance quantitative trait loci (QTL) *pi21* for resistance to *Magnaporthe oryzae*, causal agent of rice blast (Fukuoka *et al.* 2009). This diverse set of genes underlying disease QTL supports the notion that a range of genes and mechanisms are involved in QDR (Poland *et al.* 2009).

Northern leaf blight (NLB) is an endemic disease of maize throughout the world. Caused by a semi-biotrophic fungal pathogen, *Setosphaeria turcica* (anamorph *Exserohilum turcicum*), NLB is found in cool temperate climates and tropical highlands where it causes moderate to severe yield losses. Previous research on resistance to NLB has resulted in identification of multiple disease QTL on every chromosome except 10, indicating a diverse genetic architecture (Wisser *et al.* 2006). Three major genes conferring race-specific, though not complete, resistance have been mapped for *S. turcica*; *Ht1* located in maize bin 2.08, and *Ht2* and *Htn1* located in maize bin 8.06 (Wisser *et al.* 2006). Based on the economic importance of resistance to NLB and the genetic resources available in maize, namely a large reference design population, we utilized this pathosystem as a model for understanding the genetic architecture of QDR in plants.

The maize nested association mapping (NAM) population has been developed as a community resource for dissection of complex traits in maize (Yu *et al.* 2008; Buckler *et al.* 2009; McMullen *et al.* 2009). Constructed in a reference design, NAM consists of 25 recombinant inbred line (RIL) families. Each family is comprised of 200 RILs derived from a cross between one of 25 diverse founder inbred lines and the reference inbred B73. The reference design narrows the physiological maturity range which is beneficial for disease resistance evaluations. RILs from the intermated B73 x Mo17 population (IBM) were included as a 26th family (Lee *et al.* 2002). The RILs are genotyped

with 1106 SNP markers (McMullen *et al.* 2009) and are publicly available (http://www.panzea.org/lit/germplasm.html).

RESULTS

We evaluated 4,630 RILs, along with each of the 26 parents, over three seasons for resistance to NLB. Each year, lines were artificially inoculated and visually evaluated at three time-points for disease severity based on the percentage of the total leaf area covered with necrotic lesions. A multivariate mixed model was used to account for field effects while modeling unique variance and covariance structure for each of the families at each rating. The model solution gave best linear unbiased predictions (BLUPs) for disease severity on each of the RILs at each rating and an NLB index was calculated by averaging the three BLUPs for each line.

NAM captures a large range of phenotypic diversity for NLB resistance. The NAM founders were selected from a larger panel of maize inbreds that captures most of the genetic diversity of modern maize lines (Flint-Garcia *et al.* 2005). Concurrent work in our group has shown this panel also captures a wide range of resistance phenotype for NLB (Kolkman *et al.* unpublished). The founder lines were selected from this panel to maximize molecular diversity and, as a result, have large phenotypic variation for resistance to NLB (**Figure 2.1, Table 2.1**). At the third disease rating, the NAM founders varied over a 10-fold range in percent diseased foliage, while the progeny varied over nearly 40-fold range. The BLUPs for the third rating ranged from less than 5% to almost 60% for the NAM parents and from 2.5% to 80% for lines in the NAM. The common reference parent, B73, is considered moderately susceptible with a BLUP from rating 3 of 34%.

Heritability for line BLUPs was calculated for individual families and

NAM as a whole (**Table 2.2**). To reflect the multivariate data, heritability estimates were calculated for each rating and for the NLB index. Within-family heritability averaged 0.63, 0.71, 0.68 and 0.77 for the three ratings and the NLB index, respectively. The NLB index incorporates all three ratings leading to more accurate phenotypic values and higher heritability. As NAM captures variation across diverse founder lines as well as progeny segregation within families, the heritability for NAM was higher than any of the individual families. The heritability for NLB index across NAM was 0.87 and this was the phenotype used for further study and mapping.

Joint linkage mapping captures most of the genetic variance of the founder inbreds. Joint linkage mapping was conducted by stepwise selection of marker effects. A main effect for family was included in the model and marker effects were nested within families to allow for unique allele effects in each family. Relatively maturity was slightly correlated with NLB resistance. We therefore included days to anthesis (DTA) (Buckler *et al.* 2009) as a covariate to avoid identification of maturity related QTL. Stepwise GLM selection identified 30 QTL that accounted for 77% of the variance in the BLUPs (**Figure 2.2**). Given the heritability range for NAM, these QTL explain most of the genetic variance. At these 30 loci there was an average of 8.7 QTL segregating in a given population with a range of 3 to 14. At 21 of the 30 QTL (70%), estimated allele effects were both higher and lower than the reference B73, indicating the importance of allelic variants in creating a range of diverse phenotype. This agrees with the hypothesis that rare alleles at a limited number of common loci condition most of the genetic variance rather than an infinite QTL model (Buckler *et al.* 2009). Several founder lines carried only resistance alleles relative to B73.

Figure 2.1a. Founder inbred lines of NAM capture large phenotypic variation for NLB resistance.

Representative ear leaf from each of the 25 founder inbred lines of the maize nested association mapping population (NAM), B73 the NAM reference parent, and Mo17. *Setosphaeria turcica*, the causal agent of northern leaf blight (NLB), produces large cigar shaped lesions which coalesce leading to complete blighting of the leaf in more susceptible inbred lines.

Figure 2.1b. Allele effects for each of 30 QTL identified for NLB resistance.

Estimated allele effects for each founder line at marker loci identified as quantitative trait loci for NLB resistance. Significant allele effects are shown in blue (positive) and red (negative) with the size of each bar representing the effect size.

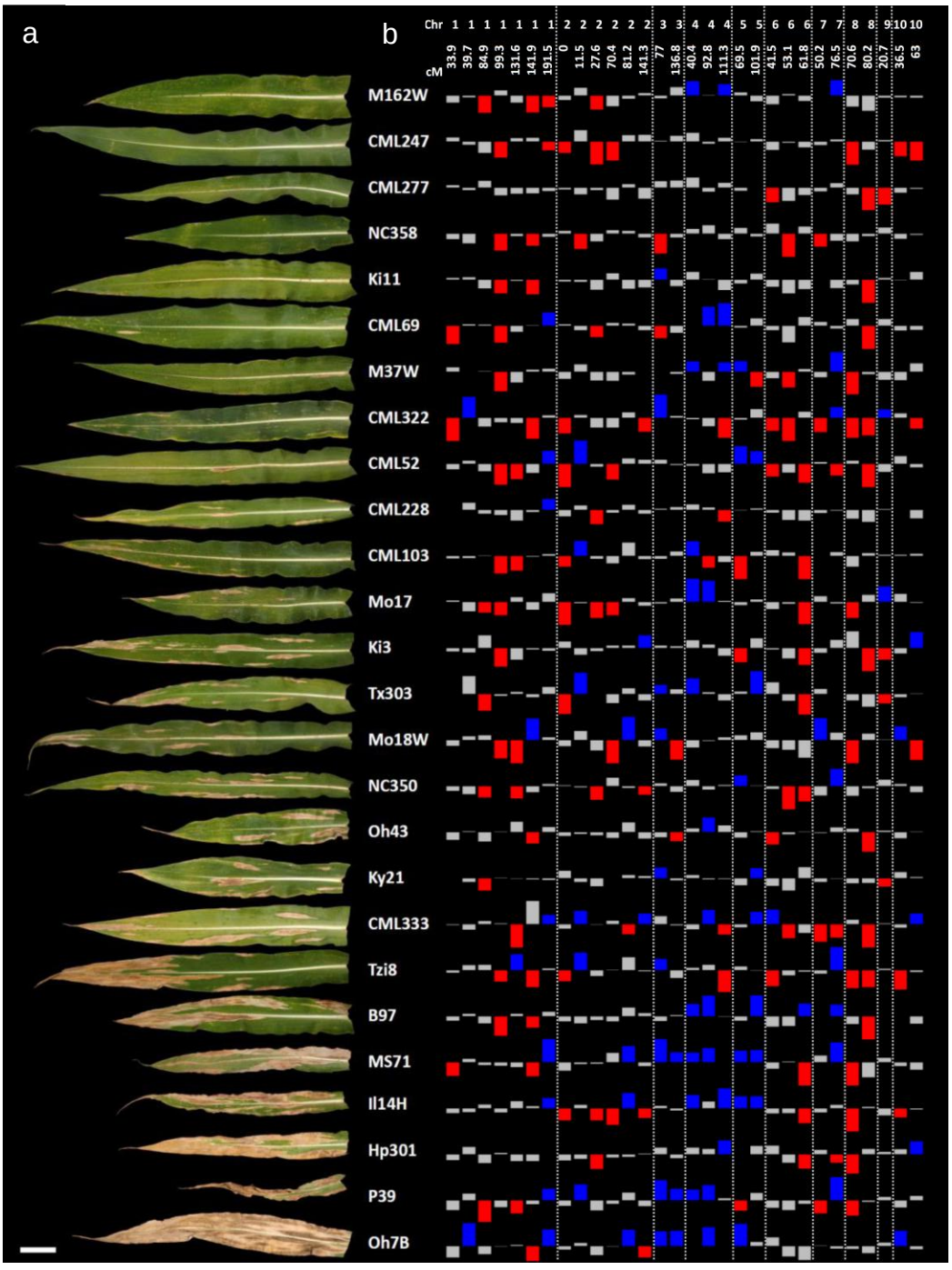

a
b
Chr 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 4 4 4 5 5 6 6 6 7 7 8 8 9 10 10
cM 33.9 39.7 84.9 99.3 131.6 141.9 191.5 0 11.5 27.6 70.4 81.2 141.3 77 136.8 40.4 92.8 111.3 69.5 101.9 41.5 53.1 61.8 50.2 76.5 70.6 80.2 20.7 36.5 63
M162W
CML247
CML277
NC358
Ki11
CML69
M37W
CML322
CML52
CML228
CML103
Mo17
Ki3
Tx303
Mo18W
NC350
Oh43
Ky21
CML333
Tzi8
B97
MS71
Il14H
Hp301
P39
Oh7B

Table 2.1 Trait summary for resistance to northern leaf blight in NAM families

Family	# of lines	Rating 1		Rating 2		Rating 3	
Founders	27	7.8	(1.3; 27.6)	11.8	(1.7; 43)	20.1	(4.1; 59.4)
B97 x B73	189	14.9	(4.1; 41.8)	22.5	(7.1; 58)	35.3	(13; 79.6)
CML103 x B73	177	9.3	(1.8; 23.4)	14.3	(2.5; 35.9)	23.5	(8.6; 45.9)
CML228 x B73	161	7.5	(2; 20.1)	11.7	(2.9; 24.9)	22.6	(8.9; 46.4)
CML247 x B73	182	7.6	(1.5; 20)	11.8	(1.6; 33.8)	23.0	(5.7; 53.3)
CML277 x B73	151	5.6	(0.6; 13.9)	8.1	(1; 20.9)	15.7	(2.7; 36.5)
CML322 x B73	175	11.8	(2.4; 30.9)	17.7	(3; 43.7)	30.0	(9.4; 62.6)
CML333 x B73	183	14.5	(1.5; 34.6)	21.6	(2.9; 50.7)	32.4	(9.5; 63.1)
CML52 x B73	163	7.6	(0.7; 21.1)	11.1	(1.5; 33.3)	19.5	(4.4; 51.4)
CML69 x B73	186	6.7	(0.7; 26.9)	10.3	(1.2; 43.7)	18.3	(3.6; 59.8)
Hp301 x B73	182	20.1	(8.2; 46)	30.4	(13; 63.4)	45.3	(27.7; 76)
Il14H x B73	187	14.1	(4.1; 31.6)	21.5	(5; 49.3)	36.1	(10.3; 75.4)
Ki11 x B73	168	9.0	(0.3; 27.2)	13.5	(0.3; 39.9)	23.4	(2.6; 52.5)
Ki3 x B73	191	6.0	(0.8; 15.5)	8.8	(1.2; 21.7)	16.4	(4.2; 34.7)
Ky21 x B73	191	16.5	(5.2; 31.5)	24.4	(10.5; 40.5)	37.8	(17.2; 60.6)
M162W x B73	187	6.5	(1; 17.4)	9.6	(0.8; 28.9)	17.5	(2.5; 43.8)
M37W x B73	187	7.0	(1.6; 22.1)	10.5	(1.9; 36.3)	19.1	(6.1; 53.2)
Mo17 x B73	179	11.8	(1.5; 32.1)	17.8	(3.5; 43.3)	29.8	(10.2; 60)
Mo18W x B73	173	9.5	(0.6; 31.5)	14.8	(1.1; 45.7)	26.1	(6.3; 60.8)
MS71 x B73	191	15.7	(2.2; 50.2)	23.4	(5.1; 62.1)	35.0	(11.1; 79.2)
NC350 x B73	182	12.3	(1.5; 28.2)	18.6	(2.3; 40.6)	29.4	(5.6; 57.4)
NC358 x B73	182	7.8	(2.2; 21.6)	12.2	(2.8; 36)	22.8	(7.5; 54.7)
Oh43 x B73	180	13.6	(6.3; 26.6)	20.7	(9.5; 40.1)	32.4	(18.3; 57.6)
Oh7B x B73	176	22.5	(6.4; 53.4)	34.5	(8.4; 73.7)	48.9	(18.3; 79.3)
P39 x B73	174	16.1	(4.5; 37.4)	24.6	(8.9; 57.4)	39.0	(17.9; 73.4)
Tx303 x B73	188	12.6	(2.3; 24.5)	19.4	(4.1; 37.1)	32.4	(12.3; 52.8)
Tzi8 x B73	182	12.7	(1.4; 41.6)	18.9	(3; 52.1)	30.9	(7.1; 73.7)
NAM	4667	11.4	(0.3; 53.4)	17.2	(0.3; 73.7)	28.2	(2.5; 79.6)
		mean	(min;max)				

Table 2.2. Heritability for resistance to northern leaf blight in NAM

Family	Rating 1	Rating 2	Rating 3	Index
NAM	0.75	0.82	0.81	0.87
B97	0.65	0.70	0.71	0.78
CML103	0.62	0.72	0.63	0.75
CML228	0.53	0.67	0.62	0.70
CML247	0.61	0.74	0.75	0.79
CML277	0.57	0.66	0.71	0.75
CML322	0.72	0.79	0.76	0.83
CML333	0.63	0.70	0.63	0.75
CML52	0.71	0.78	0.78	0.83
CML69	0.66	0.76	0.75	0.81
Hp301	0.63	0.70	0.55	0.74
Il14H	0.60	0.72	0.72	0.77
Ki11	0.74	0.81	0.78	0.85
Ki3	0.55	0.64	0.61	0.71
Ky21	0.46	0.41	0.48	0.58
M162W	0.54	0.71	0.72	0.76
M37W	0.54	0.68	0.66	0.74
Mo17	0.68	0.71	0.69	0.78
Mo18W	0.72	0.80	0.76	0.83
MS71	0.77	0.80	0.79	0.85
NC350	0.68	0.74	0.75	0.80
NC358	0.50	0.69	0.69	0.73
Oh43	0.46	0.56	0.55	0.65
Oh7B	0.69	0.78	0.65	0.79
P39	0.65	0.71	0.67	0.77
Tx303	0.64	0.71	0.64	0.76
Tzi8	0.74	0.75	0.76	0.83
Avg. within Family	0.63	0.71	0.68	0.77
Founders	0.80	0.87	0.88	0.90

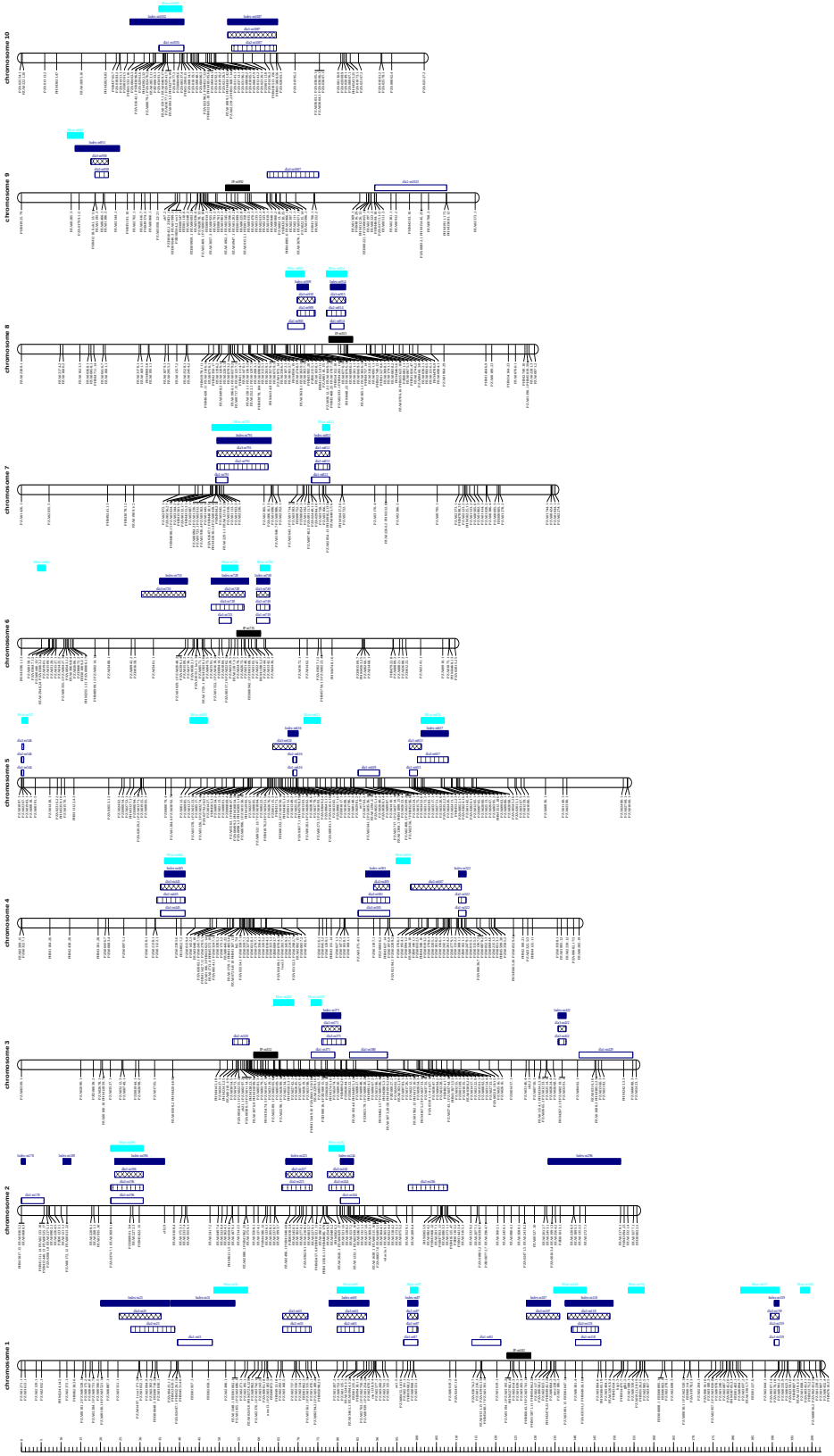
However, even the most resistant genotypes only incorporated favorable alleles at five to eight of the QTL. A hypothetical genotype incorporating the favorable allele at all 30 QTL has a predicted phenotype with negative diseased leaf area. Under very high disease pressure this 'super-genotype' would presumably remain at low or no disease severity. Phenotypic selection in an environment with moderate disease pressure would reach a phenotypic minimum, but additional progress presumably be achieved through marker assisted pyramiding of further resistance alleles.

Incubation period (IP) is a resistance component for NLB measured as the number of days past inoculation until the appearance of disease symptoms. It was hypothesized that IP could identify resistance that is effective during the early stages of pathogenesis. We modeled IP according with a Cox proportional hazards model, as IP is analogous of survival data from clinical trials. Using stepwise model selection, five QTL for IP were identified in NAM (**Figure 2.2**). The three most significant QTL for IP were consistent with large effect QTL for disease severity (Chrs. 1, 6 and 8). The two remaining QTL for IP were identified at locations not associated with disease severity. These two QTL could be effective during initial pathogen infection and establishment but not during the later stages of pathogen development.

QTL identified here were largely located in regions previously associated with QDR to NLB (Wisser *et al.* 2006). However, presumably due to greater statistical power in NAM as well as the ability to survey a broad range of germplasm, additional QTL were identified (e.g. Chr. 10). Two large effect QTL were located on Chr. 8 (137.5-Mb and 152.2-Mb). This location is consistent with the race-specific resistance gene *Ht2* (C. Chung *et al.* submitted).

Figure 2.2. NAM consensus genetic map and identified QTL for NLB.

IP –black; Rating 1 – open; Rating 2 – stripe; Rating 3- checkered; NLB index – filled dark; Multivariate – filled light


chromosome 10
chromosome 9
chromosome 8
chromosome 7
chromosome 6
chromosome 5
chromosome 4
chromosome 3
chromosome 2
chromosome 1

Though race-specific, *Ht2* does not condition complete immunity and could be considered a large effect QTL. This was consistent with our observation, as no population had a bimodal distribution or other phenotypic evidence of a "major" gene.

To assess the proportion of genetic variance captured in the QTL identified, the sum of all QTL effects from the additive QTL model for each founder line was compared to the actual phenotype of the parents (**Figure 2.3**). There was a strong and significant correlation between the predicted founder phenotypes and the empirical observations. The coefficient of determination was 0.60 for the regression of the BLUP phenotype on the sum of significant allele effects (p-value<0.05). The correlation between the founder inbred BLUPs and the sum of QTL effects was higher when all allele effects were included (R^2 = 0.71), indicating that very small 'non-significant' allele effects at QTL contribute to the resistance phenotype. The heritability of NLB index for the founders of NAM was 0.74. Considering only 'significant' QTL alleles, the additive model explains 80% of the heritable genetic variation in the NAM founders. When the 'non-significant' allele effects are included, the R^2 increased to 0.71, approaching the heritability for this trait, explaining 96% of the genetic variance for NLB resistance in the founders of NAM. The remaining genetic variance not explained by this simple additive model is likely due to a combination of higher order epistatic interactions and a lack of power to detect extremely small QTL.

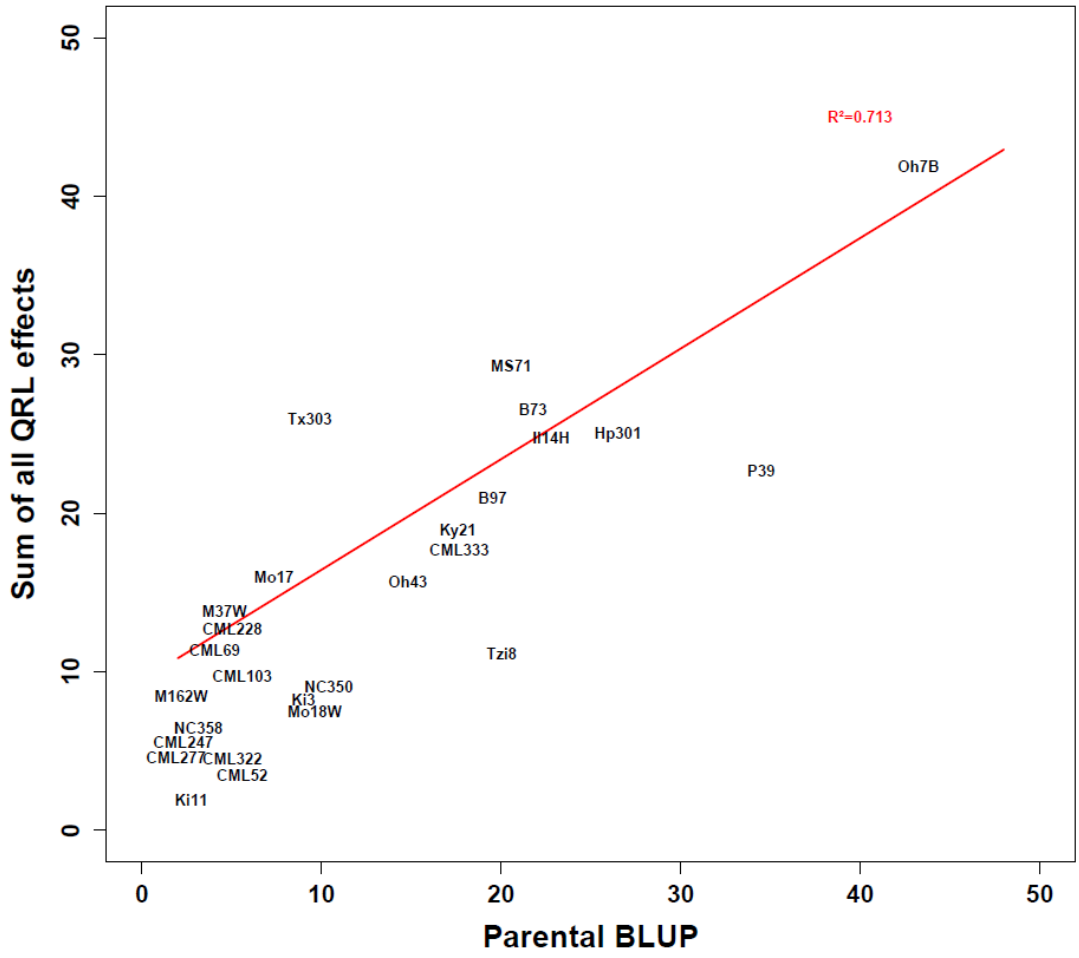


Figure 2.3 Additive genetic model for NLB resistance gives accurate phenotypic prediction of NAM founders.

Linear fit of predicted phenotypic values for each of the diverse founder inbred lines of NAM against the empirical phenotypic observations. Units are shown in percentage diseased leaf area.

Most of the QTL had very small estimated effects. Of the 200 QTL alleles identified (p-value < 0.05), only 24 conditioned an increase or decrease of more than 3% of blighted leaf area. This is about the lower limit that can be visually distinguished by a trained scorer. Only three QTL alleles had an estimated effect on diseased leaf area > 5%. Most of the QTL allele effects were estimated as too small to visually distinguish without statistical inference. This is consistent with results from flowering time in maize for which most alleles had effect sizes less than the observation unit of one day (Buckler *et al.* 2009). Every founder inbred carried both positive and negative alleles. Lines such as M162W and CML247 that had near immunity carried alleles for susceptibility, and Oh7B that is almost completely blighted by the end of the growing season had several alleles for resistance.

Connecting defense related genes with QDR through nested association mapping. Genomewide association testing was evaluated for 1,606,526 SNPs on the NAM founders from the first generation maize haplotype map (**Figures 2.4**). Due to the low linkage disequilibrium (LD) in maize, 1.6 M SNPs were expected cover about 30-40% of the genome with LD sufficient for association testing. SNP genotypes were imputed from founder lines to RIL progeny using pedigree and linkage marker information. Trait-marker associations were then evaluated by chromosome. For each chromosome, a linear model was fit for the imputed SNPs against the residual phenotypic values for that chromosome, after accounting for all QTL on other chromosomes. At each of the QTL positions identified from joint linkage mapping, the 5 most significant SNP markers within the QTL confidence

interval were evaluated as polymorphisms within or close (<10-kb) to candidate genes (Table 2.3). Nine additional regions that appeared to contain QTL from the genome profile of SNP association results, but which were not identified as QTL, were also examined for candidate genes from the SNP association.

It has been hypothesized that the genes conditioning QDR will cover a broad range of mechanisms, including classes of genes previously unassociated with disease resistance (Poland *et al.* 2009). This hypothesis has been supported with the identification of three novel disease resistance genes underlying QTL (Fu *et al.* 2009; Fukuoka *et al.* 2009; Krattinger *et al.* 2009). As little is known about the molecular basis of QDR, it is difficult to link many of the candidate gene associations identified with disease resistance. At the same time, it is difficult to exclude any candidate gene association for the reasons outlined above. Further experience has shown that the causative polymorphism for a QTL can be distant from the functional gene (Salvi *et al.* 2007). While recognizing these limitations, we were able to link several previously know pathogen defense related genes to QDR using genome-wide nested association mapping.

Figure 2.4 Profile of SNP associations tested for resistance to northern leaf blight.

Using the NAM reference design, 1,606,526 SNPs were imputed from the NAM founders and tested for association with resistance to northern leaf blight. Each circle represents a tested SNP plotted at the genomic position of the respective SNP and log (1/prob.) from the association test. Regions within confidence intervals of QTL are shown in red.

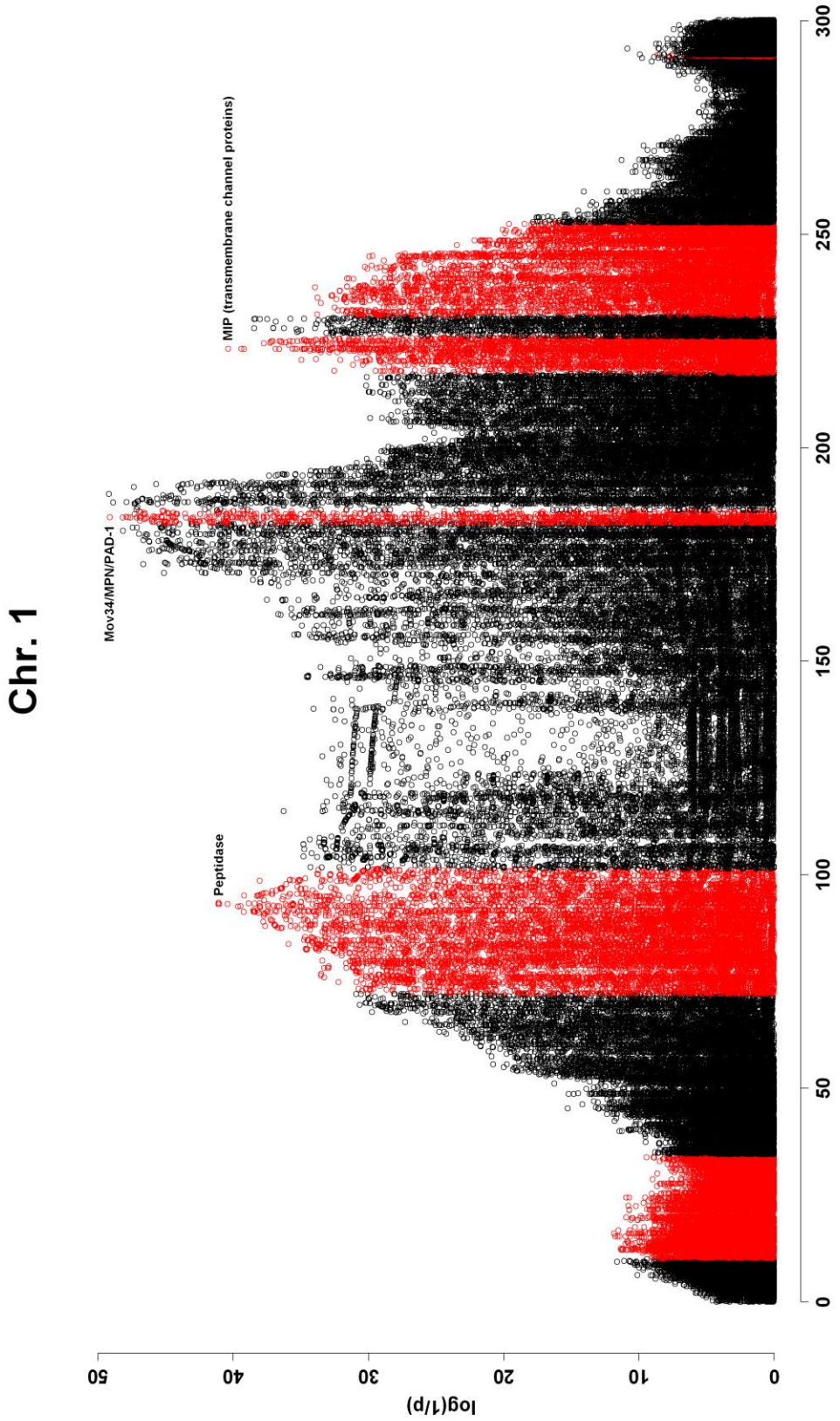
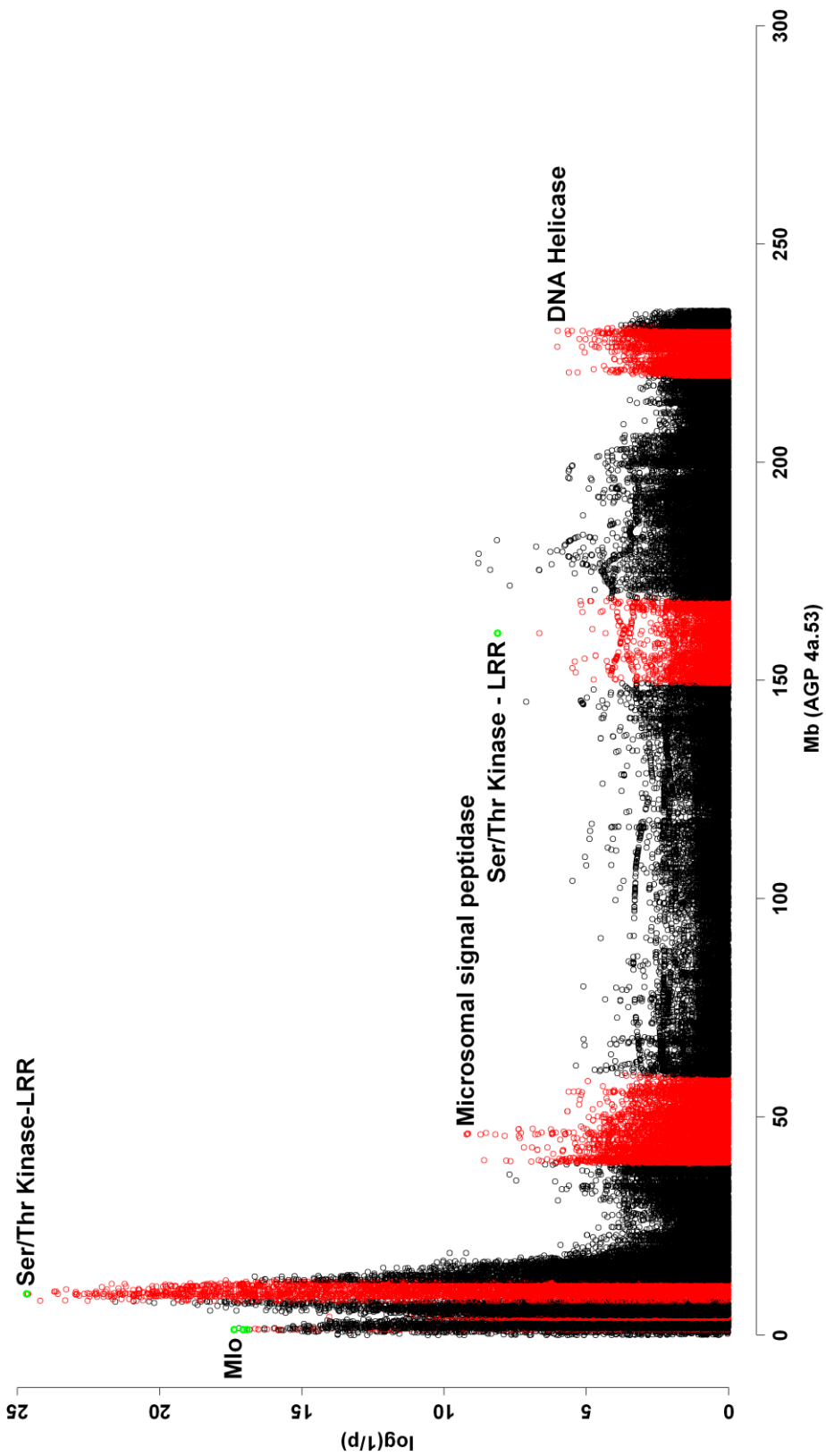


Figure 2.4 Profile of SNP associations tested for resistance to northern leaf blight

Figure 2.4 (Continued)



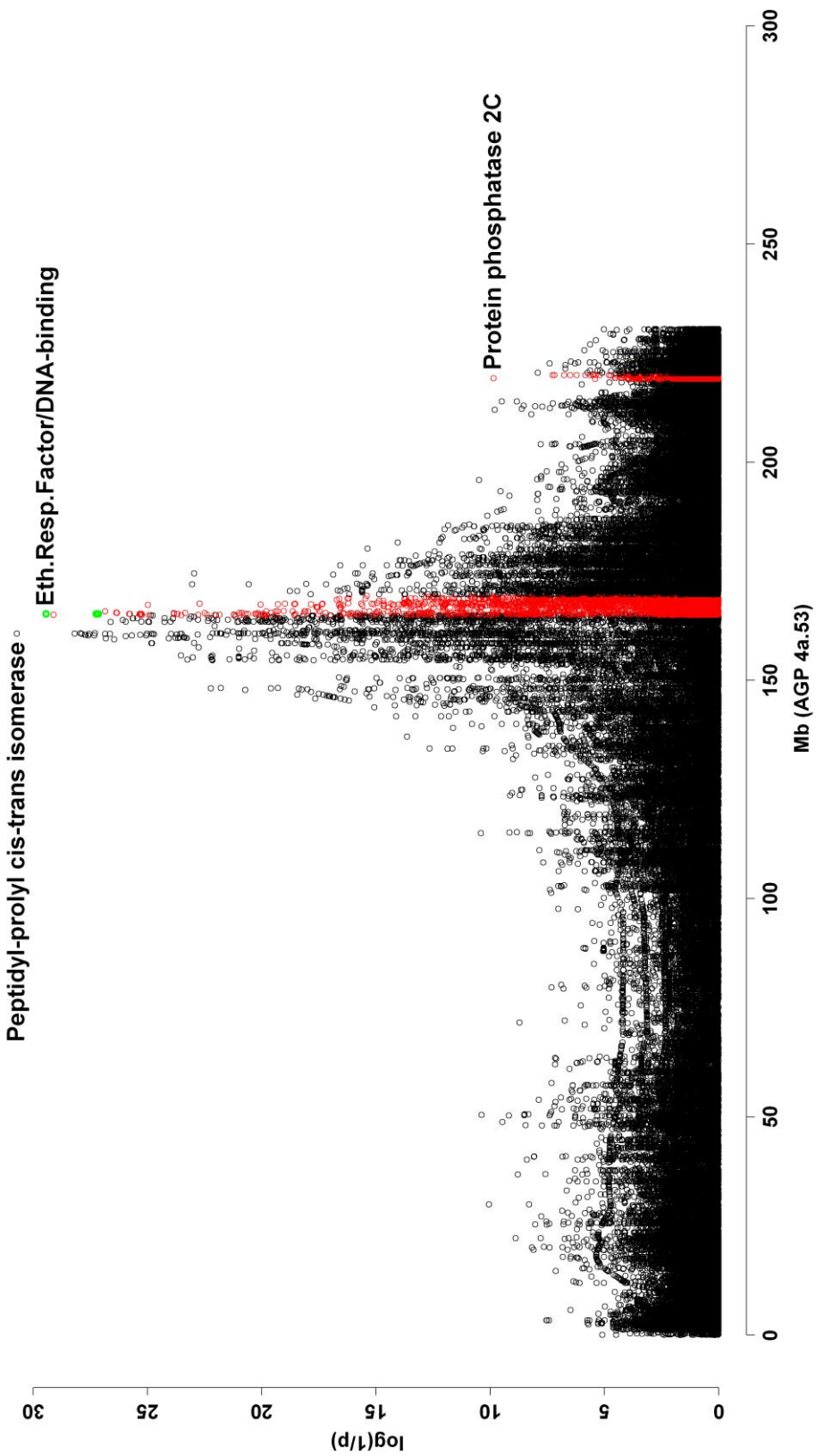


Figure 2.4 (Continued)

Figure 2.4 (Continued)

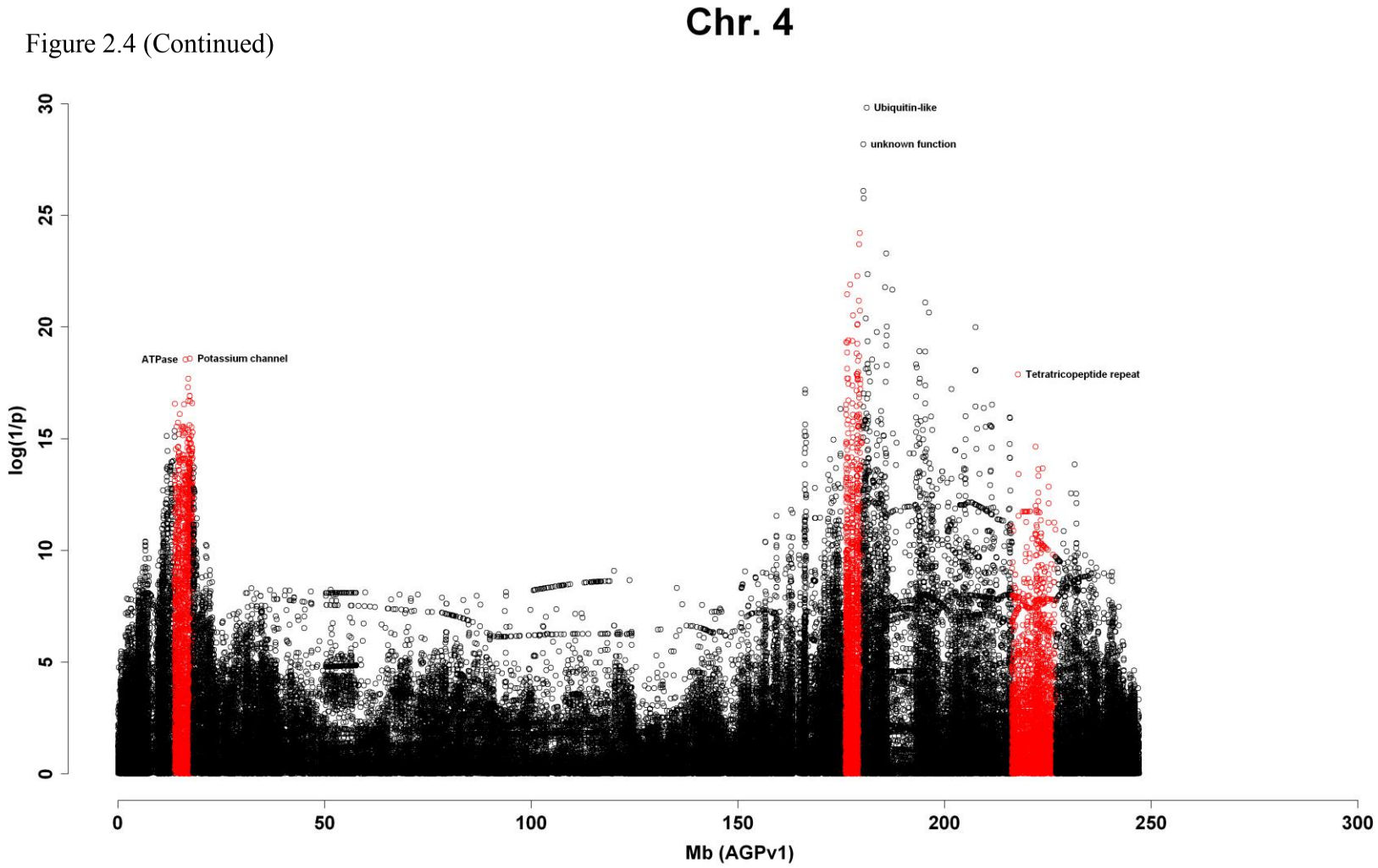


Figure 2.4 (Continued)

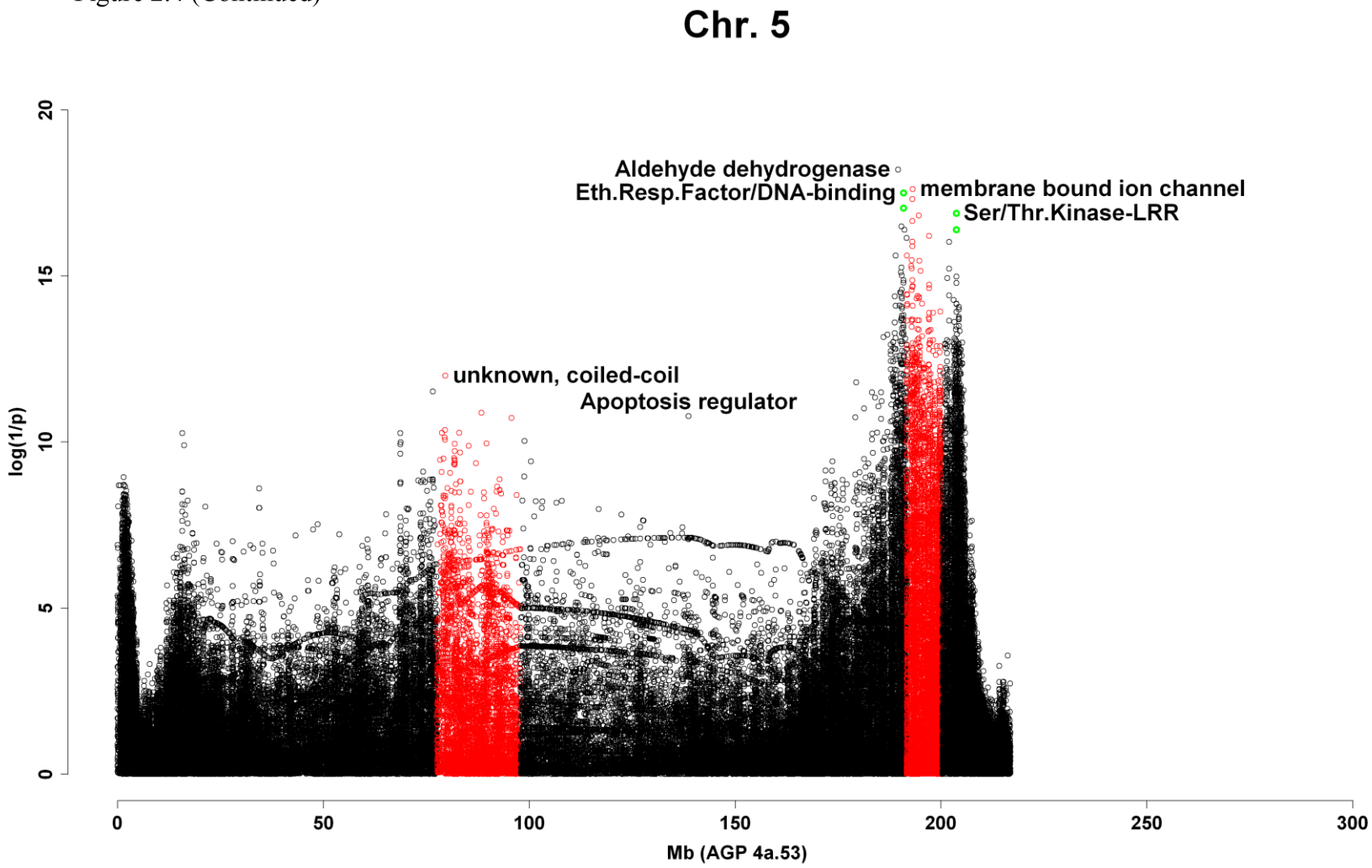


Figure 2.4 (Continued)

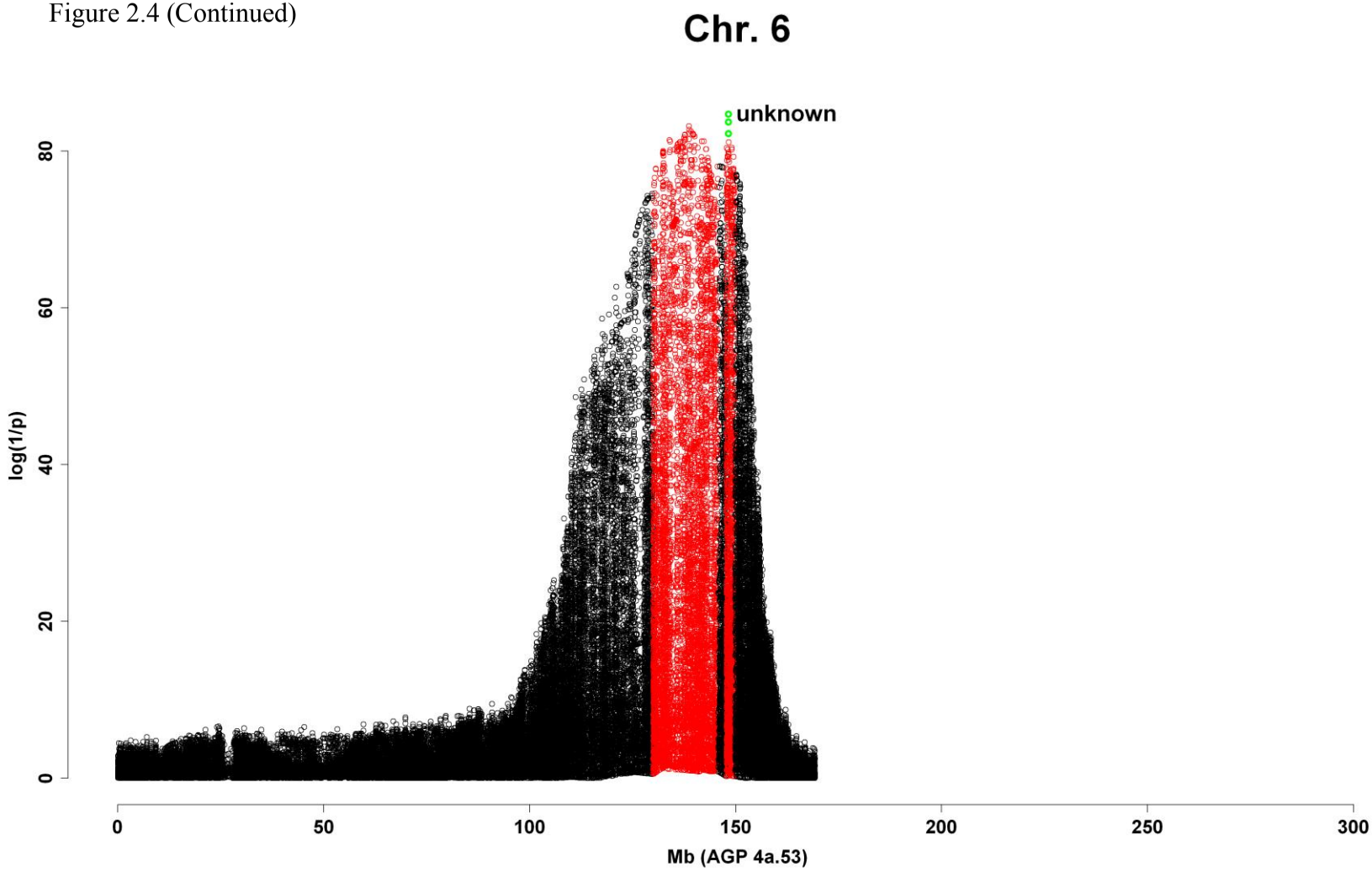


Figure 2.4 (Continued)

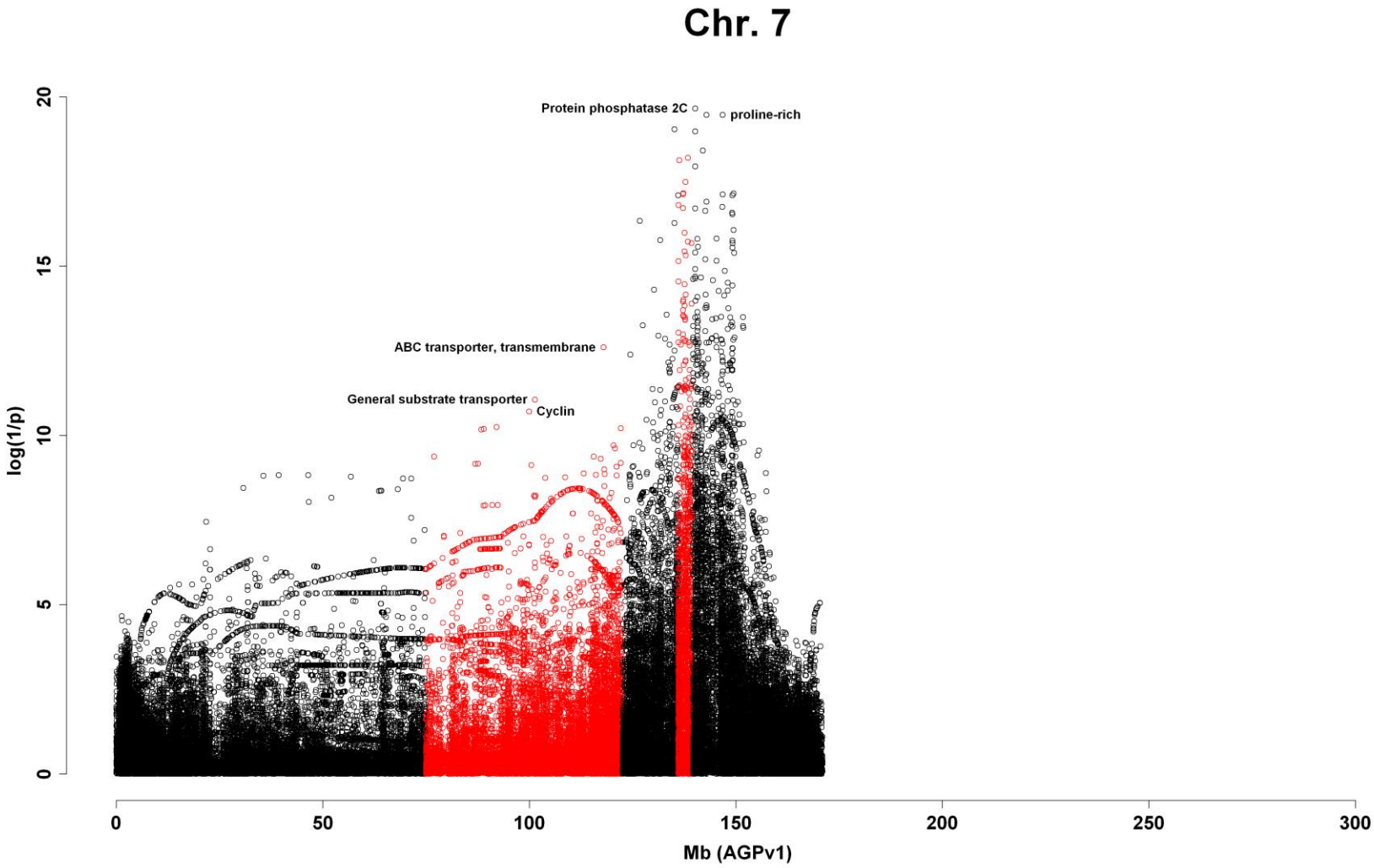


Figure 2.4 (Continued)

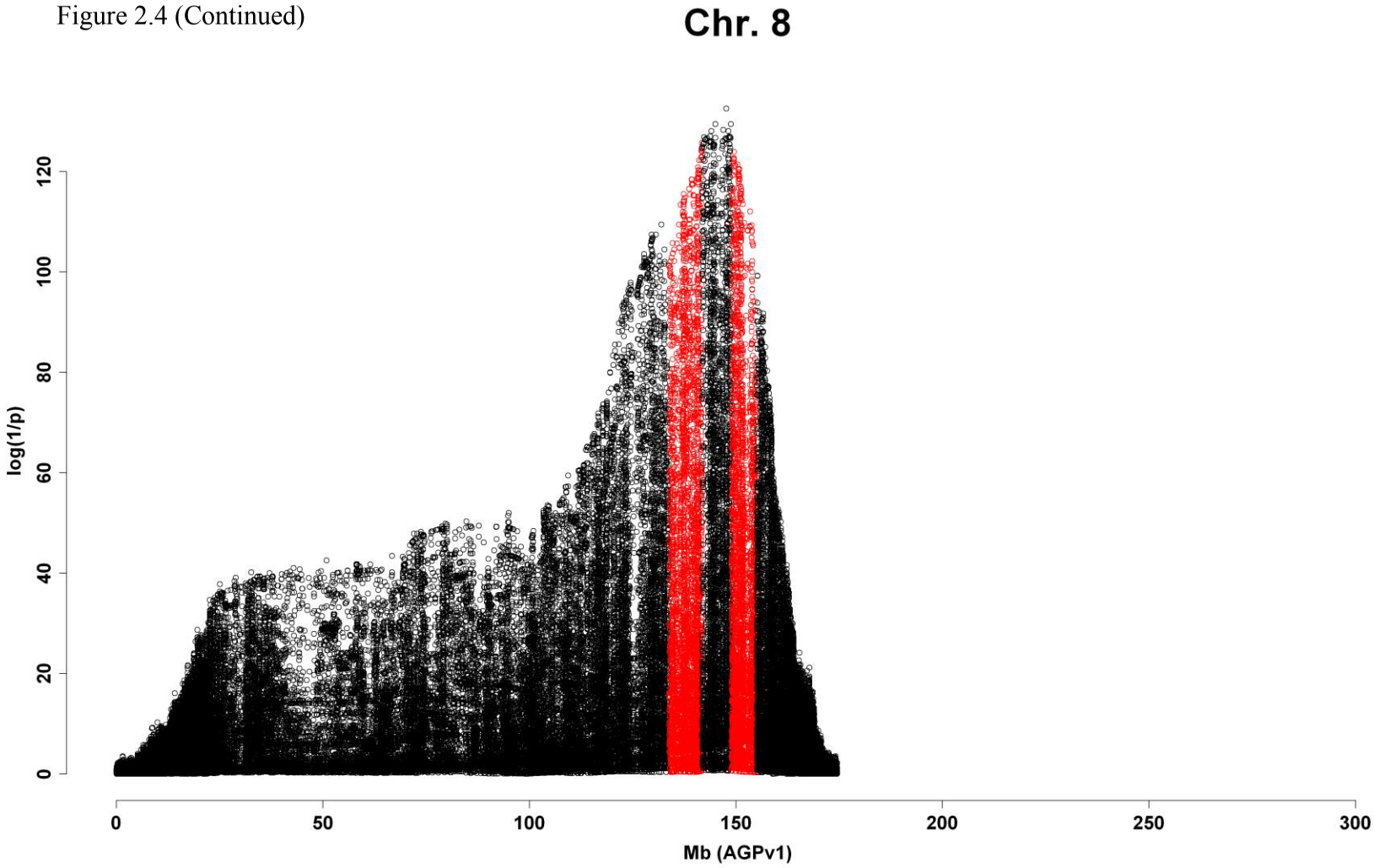


Figure 2.4 (Continued)

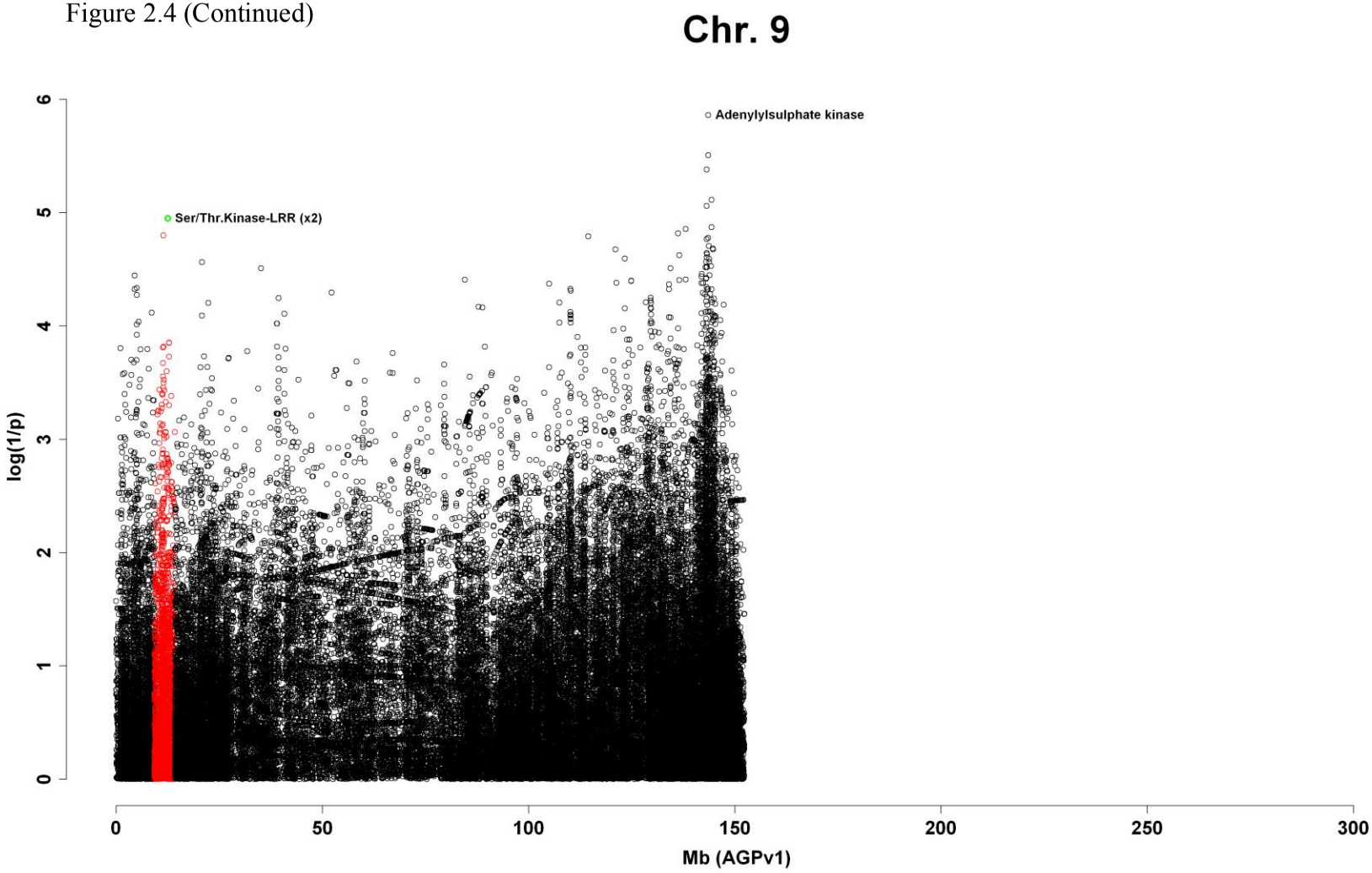


Figure 2.4 (Continued)

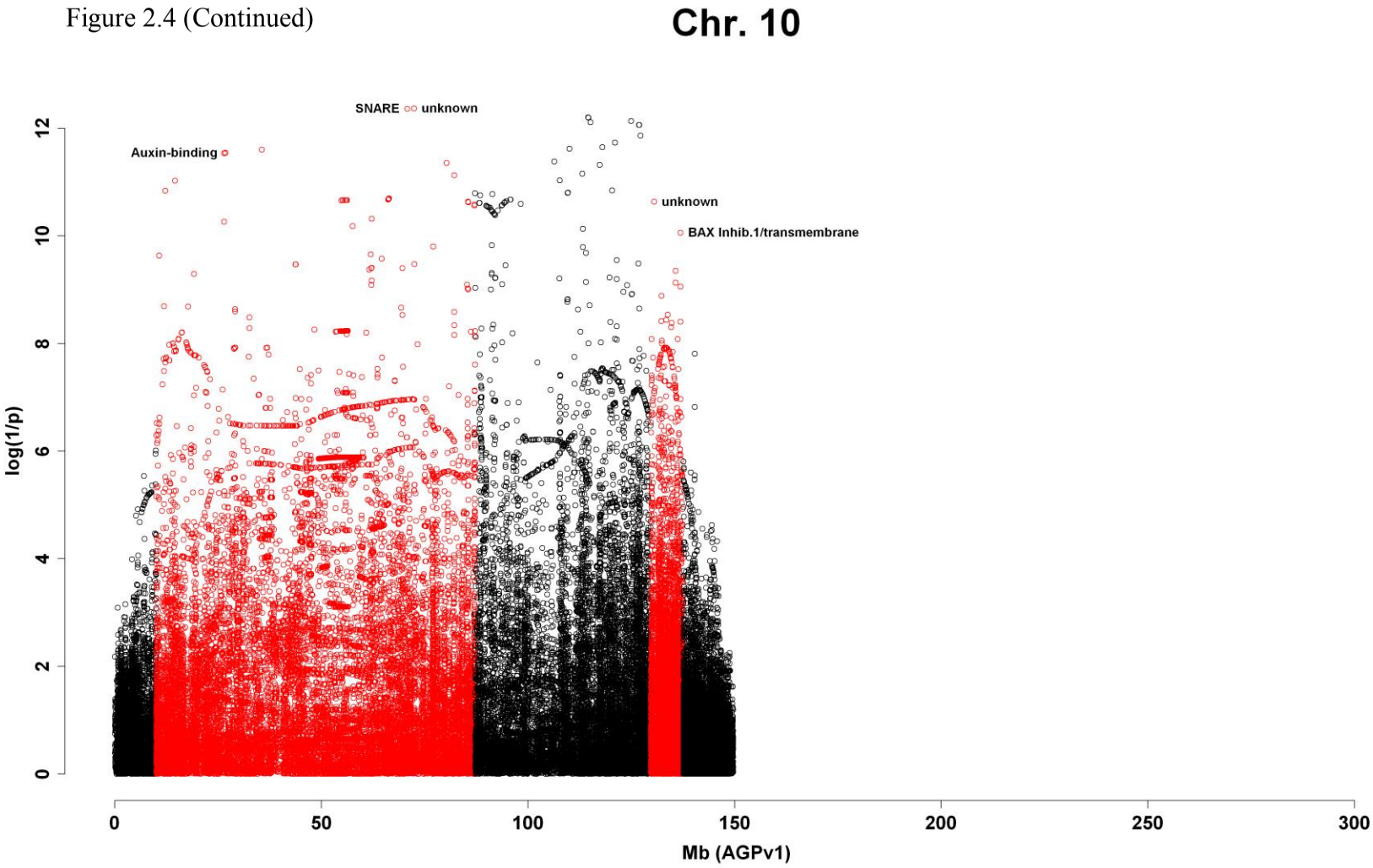


Table 2.3 Single nucleotide polymorphisms and associated candidate genes identified with genome-wide nested association mapping

Table 2.3. Single nucleotide polymorphisms and associated candidate genes identified with NAM

QRL condfidence intervals:

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
1	16009569| 16009574	11.85| 11.85	GRMZM2G178852	15997196	15998530	-1	IPR000104	Antifreeze protein, type I
			GRMZM2G703504	16014562	16019143	-1	na	unknown
1	15304297| 15304357	11.76| 11.57	GRMZM2G124353	15303636	15307028	1	IPR005814| IPR015424|	Aminotransferase class-III| Pyridoxal phosphate-dependent transferase, major region|
1	16746663	11.54| 11.54	GRMZM2G445634	16741838	16742760	1	IPR010399| IPR018467|	Tify| CCT domain-like|
			GRMZM2G445630	16746105	16746859	1	na	unknown
			GRMZM2G147920	16748343	16750546	1	IPR001578	Peptidase C12, ubiquitin carboxyl-termil hydrolase 1
			GRMZM2G147910	16751666	16752780	1	na	unknown
1	24392170| 24394831	10.95| 10.72	GRMZM2G139852	24319353	24323290	-1	IPR006702	Uncharacterised protein family UPF0497, trans-membrane plant
1	22489456	10.78	GRMZM2G077127	22463288	22465781	-1	IPR000073	Alpha/beta hydrolase fold-1
1	21533708| 21533734	10.71| 10.71	GRMZM2G086604	21514408	21532752	-1	IPR006702	Uncharacterised protein family UPF0497, trans-membrane plant
1	93187932| 93189171| 93189147	41.12| 41.05| 40.41	GRMZM2G033478	93032295	93035782	1	IPR008395| IPR014002|	Agenet| Tudor-like, plant|
1	93444878	41.08	GRMZM2G113098	93377448	93379143	1	IPR000679| IPR002062| IPR005398| IPR016679|	Zinc finger, GATA-type| Oxytocin receptor| Tubby, N-termil| Transcription factor, GATA, plant|
1	**92531926**	**39.64| 39.64**	**GRMZM2G104016**	**92529129**	**92530269**	**1**	**IPR001471| IPR016177|**	**Pathogenesis-related transcriptiol factor and ERF, D-binding| D-binding, integrase-type|**
1	183754852| 183755280| 183755812| 183755846| 183755953	49.13| 49.13| 49.13| 49.13| 49.13	GRMZM2G080652	183753967	183757883	-1	IPR000555	Mov34/MPN/PAD-1
1	223216297| 223215875| 223216654	40.38| 40.38| 39.4	GRMZM2G303132	223212197	223213222	-1	na	unknown
1	223161205	39.19	GRMZM2G181231	223151182	223153172	1	IPR001245| IPR002290| IPR017441	Tyrosine protein kise| Serine/threonine protein kise| Protein kise, ATP binding site
1	224495114	37.73| 37.73	GRMZM2G438386	224493494	224497305	-1	IPR001117| IPR008972|	Multicopper oxidase, type 1| Cupredoxin|
1	237635240	33.96	GRMZM2G118917	237621005	237640428	1	IPR002121| IPR002562| IPR012337	Helicase and Rse D C-termil, HRDC| 3'-5' exonuclease| Polynucleotidyl transferase, Ribonuclease H fold
1	231242248	33.8	AC233892.1_FG001	231205820	231207568	-1	IPR002885	Pentatricopeptide repeat
1	235161009	33.77	GRMZM2G118173	235103801	235104758	-1	IPR000217| IPR000363| IPR000817| IPR003536| IPR008119| IPR017975|	Tubulin| Adrenergic receptor, alpha-1A| Prion protein| Translocated intimin receptor| Dense granule Gra6 protein| Tubulin, conserved site|
1	233923866	33.52	GRMZM2G069427	233904715	233907617	-1	IPR000235	Ribosomal protein S7
1	235578884	32.73| 32.73	GRMZM2G301031	235569645	235576621	-1	IPR001478	PDZ/DHR/GLGF
			GRMZM2G001606	235577013	235579062	1	IPR000381| IPR014977|	Inhibin, beta B subunit| WRC|
			GRMZM2G001592	235582346	235584261	1	IPR014977	WRC
			GRMZM2G001577	235586709	235587926	-1	IPR014977	WRC
1	292081546	8.632	GRMZM2G061900	292077150	292080731	1	IPR001806| IPR002041| IPR002078| IPR003579| IPR005225| IPR006688| IPR013684| IPR013753|	Ras GTPase| Ran GTPase| R polymerase sigma factor 54, interaction| Ras small GTPase, Rab type| Small GTP-binding protein| ADP-ribosylation factor| Miro-like| Ras|
		8.632	GRMZM2G061897	292081220	292081672	-1	IPR015880	Zinc finger, C2H2-like
1	291705178	7.717	GRMZM2G405690	291647532	291651048	-1	IPR011046	WD40 repeat-like
1	291825704| 291825654	7.513| 7.513	AC225147.4_FG003	291817648	291821080	-1	IPR011011	Zinc finger, FYVE/PHD-type
			AC225147.4_FG002	291821868	291822758	-1	IPR005679| IPR016027|	Ribosomal protein S12, bacterial-type| Nucleic acid-binding, OB-fold-like|
		7.513| 7.513	**AC225147.4_FG004**	**291832078**	**291835577**	**1**	**IPR001471| IPR016177|**	**Pathogenesis-related transcriptiol factor and ERF, D-binding| D-binding, integrase-type|**
1	291952682	6.825| 6.825	GRMZM2G155580	291930834	291939627	1	IPR015211| IPR016024|	Peptidase M1, leukotriene A4 hydrolase, aminopeptidase C-termil| Armadillo-type fold|
2	**1224121| 1227845**	**16.92| 16.53**	**GRMZM2G051974**	**1214080**	**1219617**	**1**	**IPR004326**	**Mlo-related protein**
			GRMZM2G051917	1220474	1225731	-1	IPR007603	unknown function DUF580

Table 2.3 (Continued)

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
2	1422733	16.64	GRMZM2G018595	1422634	1426449	1	IPR001245| IPR002290| IPR017441| IPR017969|	Tyrosine protein kise| Serine/threonine protein kise| Protein kise, ATP binding site| Heavy-metal-associated, conserved site|
			GRMZM2G019500	1424335	1428966	-1	na	unknown
2	1282096	16.53	GRMZM2G051917	1220474	1225731	-1	IPR007603	unknown function DUF580
2	1341314	16.22| 16.22	GRMZM2G155877	1338802	1340491	1	na	unknown
			GRMZM2G457621	1346707	1348278	1	na	unknown
2	4211378	14.03	GRMZM2G036297	4209163	4212840	1	IPR004333	Transcription factor, SBP-box
2	4313668| 4313967	11.23| 10.98	GRMZM2G019404	4306207	4313181	-1	IPR004014| IPR005834|	ATPase, P-type cation-transporter, N-termil| Haloacid dehalogese-like hydrolase|
			GRMZM2G020043	4307843	4309885	1	na	unknown
2	3851606| 3846963	10.57| 10.51	GRMZM2G087201	3846517	3851192	1	IPR004147| IPR011009|	ABC-1| Protein kise-like|
2	9394756| 9387079	24.68| 24.61	GRMZM2G380518	9389392	9390474	-1	IPR003071| IPR004064|	Orphan nuclear receptor, HMR type| EDG-6 sphingosine 1-phosphate receptor|
			GRMZM2G080041	**9395569**	**9399680**	**1**	**IPR001245| IPR001611| IPR002290| IPR011009|**	**Tyrosine protein kise| Leucine-rich repeat| Serine/threonine protein kise| Protein kise-like|**
			GRMZM2G080054	9400025	9401900	-1	IPR001092| IPR011598|	Basic helix-loop-helix dimerisation region bHLH| Helix-loop-helix D-binding|
2	7907073	24.19	GRMZM2G454511	7898380	7902637	1	IPR000858| IPR001245| IPR001480| IPR002290| IPR003609| IPR013227| IPR017441	S-locus glycoprotein| Tyrosine protein kise| Curculin-like (mannose-binding) lectin| Serine/threonine protein kise| Apple-like| PAN-like, type 2| Protein kise, ATP binding site
2	9930592| 9926025	23.70| 23.69	GRMZM2G007914	9920220	9923735	1	IPR003822| IPR004226|	Paired amphipathic helix| Tubulin binding cofactor A|
			GRMZM2G007791	9924591	9932115	-1	IPR000817| IPR003100| IPR003165| IPR012337| IPR014811	Prion protein| Argoute and Dicer protein, PAZ| Stem cell self-renewal protein Piwi| Polynucleotidyl transferase, Ribonuclease H fold| Region of unknown function DUF1785
			GRMZM2G007713	9936947	9938056	-1	IPR006121	Heavy metal transport/detoxification protein
2	45968986| 45969047	9.219| 9.219	GRMZM2G131448	45946632	45949066	-1	na	unknown
2	46100912	9.179	GRMZM2G143891	46028259	46029495	1	na	unknown
2	46124883	9.172	GRMZM2G143891	46028259	46029495	1	na	unknown
2	46136486	9.168| 9.168	GRMZM2G143891	46028259	46029495	1	na	unknown
2	**160834095| 160833488| 160833893| 160834109| 160834447**	**8.113| 6.651| 6.651| 6.651| 6.651**	**GRMZM2G163724**	**160832461**	**160835729**	**-1**	**IPR001245| IPR001611| IPR002290| IPR011009|**	**Tyrosine protein kise| Leucine-rich repeat| Serine/threonine protein kise| Protein kise-like|**
2	226449916	6.014	GRMZM2G022763	226422367	226428811	-1	IPR000504| IPR006529|	R recognition motif, RNP-1| U2 snRNP auxilliary factor, large subunit, splicing factor|
2	230097364| 230097453	6.009| 5.658	GRMZM2G414496	230090544	230094460	-1	IPR003034| IPR005161| IPR006164| IPR007188| IPR016194	D-binding SAP| Ku70/Ku80, N-termil alpha/beta| D helicase, ATP-dependent, Ku type| Arp2/3 complex, 34kDa subunit p34-Arc| Spen Paralogue and Orthologue C-termil-like
2	229350458	5.635	GRMZM2G149935	229343923	229348478	1	IPR001079| IPR002659| IPR008985	Galectin, carbohydrate recognition domain| Glycosyl transferase, family 31| Concavalin A-like lectin/glucase
			AC196006.3_FG007	229350903	229352559	1	IPR005123| IPR006620|	2OG-Fe(II) oxygese| Prolyl 4-hydroxylase, alpha subunit|
			GRMZM2G310548	229353754	229354703	1	na	unknown
2	220572811	5.617| 5.617	GRMZM2G130379	220550618	220552170	-1	IPR004039	Rubredoxin-type Fe(Cys)4 protein
3	165210392	29.44	GRMZM2G176576	165123209	165125929	-1	IPR000104| IPR007657|	Antifreeze protein, type I| Glycosyltransferase AER61, uncharacterized|
3	165031908	29.12	GRMZM2G406951	165030183	165032256	-1	IPR000223| IPR002921|	Peptidase S26A, sig| peptidase I| Lipase, class 3|
3	**165230453| 165230084**	**27.24| 27.15**	**GRMZM2G141638**	**165226713**	**165230534**	**1**	**IPR001471| IPR016177|**	**Pathogenesis-related transcriptiol factor and ERF, D-binding| D-binding, integrase-type|**
			GRMZM2G141632	165231913	165234129	1	IPR003072| IPR006041|	Orphan nuclear receptor, NOR1 type| Pollen Ole e 1 allergen and extensin|
3	165805755	26.85| 26.85	GRMZM2G093119	165791458	165800810	1	IPR006909	Rad21/Rec8 like protein, C-termil
3	219271420	9.849	GRMZM2G134628	219269938	219275703	1	na	unknown
3	219996825	7.265	GRMZM2G087628	219990983	219993332	-1	IPR003029| IPR016027|	S1, R binding| Nucleic acid-binding, OB-fold-like|
			GRMZM2G087675	219993354	219995459	1	na	unknown
			GRMZM2G087769	219997067	219998476	1	IPR013216	Methyltransferase type 11

Table 2.3 (Continued)

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
3	219965383| 219971483| 219966168	7.18| 6.759| 6.49	GRMZM2G059392	219963168	219967307	-1	IPR001952| IPR005269|	Alkaline phosphatase| Conserved hypothetical protein CHP00730|
4	17396737| 17393544	18.59| 16.92	GRMZM2G149108	17390354	17396704	-1	IPR000594| IPR005404| IPR009036	UBA/THIF-type D/FAD binding fold| Potassium channel, voltage dependent, Kv3.3| Molybdenum cofactor biosynthesis, MoeB
4	16331322	18.54	GRMZM2G024527	16264679	16265482	1	IPR000626	Ubiquitin
			GRMZM2G315264	16331959	16339834	1	IPR003593| IPR003960| IPR011704	ATPase, AAA+ type, core| ATPase, AAA-type, conserved site| ATPase associated with various cellular activities, AAA-5
4	16991887	17.68	GRMZM2G117627	16960217	16961827	1	IPR002921	Lipase, class 3
4	16979105	17.31	GRMZM2G117627	16960217	16961827	1	IPR002921	Lipase, class 3
4	179465005	24.21	GRMZM2G041699	179456828	179458730	-1	IPR002213	UDP-glucuronosyl/UDP-glucosyltransferase
4	179289074	23.71	GRMZM2G108919	179274951	179282336	1	IPR000504| IPR002344| IPR006030	R recognition motif, RNP-1| Lupus La protein| Molluscan rhodopsin C-termil tail
			GRMZM2G109140	179285048	179288335	-1	IPR001810	Cyclin-like F-box
4	178887765	22.29	GRMZM2G008819	178879058	178880132	-1	IPR003245| IPR008972|	Plastocyanin-like| Cupredoxin|
			GRMZM2G008691	178887505	178889732	-1	IPR001092| IPR011598|	Basic helix-loop-helix dimerisation region bHLH| Helix-loop-helix D-binding|
4	177207769	21.91	GRMZM2G158860	177138235	177139052	-1	IPR001245| IPR002290| IPR017441	Tyrosine protein kise| Serine/threonine protein kise| Protein kise, ATP binding site
4	176431549	21.46| 21.46	GRMZM2G121151	176414136	176415723	1	IPR007087	Zinc finger, C2H2-type
			GRMZM2G139293	176432268	176432986	1	na	unknown
			GRMZM2G441144	176434176	176434964	-1	na	unknown
4	217776565	17.87	GRMZM2G111455	217775125	217775709	1	na	unknown
4	**222062857**	**14.64**	**GRMZM2G138896**	**221994533**	**221996601**	**1**	**IPR001938**	**Thaumatin, pathogenesis-related**
			GRMZM2G163801	222063136	222064418	1	IPR001810| IPR017451|	Cyclin-like F-box| F-box associated type 1|
4	**223761489**	**13.68**	**GRMZM2G075925**	**223696412**	**223697414**	**1**	**IPR001929| IPR006045| IPR011051| IPR013096|**	**Germin| Cupin 1| Cupin, RmlC-type| Cupin 2, conserved barrel|**
4	222720205	13.63	GRMZM2G481440	222718206	222723077	-1	IPR002528	Multi antimicrobial extrusion protein MatE
			GRMZM2G481449	222722612	222723274	1	na	unknown
4	217925324	13.42| 13.42	AC184764.3_FG003	217920305	217921156	1	IPR000104| IPR000870|	Antifreeze protein, type I| Homoserine kise|
			GRMZM2G158281	217928245	217934307	1	IPR001092| IPR011598|	Basic helix-loop-helix dimerisation region bHLH| Helix-loop-helix D-binding|
5	79584201	12	GRMZM2G024612	79573447	79584428	-1	IPR009053	Prefoldin
5	88381890	10.88	GRMZM2G417525	88289238	88295028	-1	na	unknown
5	95691862	10.72	GRMZM2G462647	95681829	95683887	1	na	unknown
5	79494232| 79494224	10.36| 10.36	GRMZM2G067313	79490307	79492993	1	IPR004140	Exo70 exocyst complex subunit
			GRMZM2G367431	79494101	79494866	1	IPR006461	unknown function Cys-rich
			GRMZM2G067350	79497397	79502243	-1	na	unknown
5	193127131	17.62	GRMZM2G465957	193040453	193042688	1	IPR002290| IPR011009|	Serine/threonine protein kise| Protein kise-like|
		17.62	GRMZM2G005996	193129129	193133348	1	IPR006685| IPR010920| IPR016688	Mechanosensitive ion channel MscS| Like-Sm ribonucleoprotein-related, core| Membrane protein, At2g17000, predicted
		17.62	GRMZM2G005909	193133640	193135071	-1	IPR000092| IPR008949|	Polyprenyl synthetase| Terpenoid synthase|
5	193030547| 193030619	17.31| 16.65	GRMZM2G165601	193027219	193028901	1	na	unknown
			GRMZM2G465953	193032452	193032847	1	na	unknown
5	194597978	16.82	GRMZM2G404617	194464754	194467080	-1	IPR003653	Peptidase C48, SUMO/Sentrin/Ubl1
5	197099274	16.21| 16.21	GRMZM2G075265	197092421	197096418	1	IPR000569| IPR004856|	HECT| Glycosyltransferase, ALG6/ALG8|
			GRMZM2G375153	197105147	197106711	1	IPR011016	Zinc finger, RING-CH-type
6	114031358	12.1	GRMZM2G165969	114015216	114031842	1	IPR001254| IPR001478| IPR009003	Peptidase S1 and S6, chymotrypsin/Hap| PDZ/DHR/GLGF| Peptidase, trypsin-like serine and cysteine
6	138675953| 138675876| 138675782	83.19| 83.19| 83.19	GRMZM2G117357	138661350	138676002	-1	IPR001753| IPR006108| IPR008927| IPR016040| IPR018376	Crotose, core| 3-hydroxyacyl-CoA dehydrogese, C-termil| 6-phosphoglucote dehydrogese, C-termil-like| D(P)-binding| Enoyl-CoA hydratase/isomerase, conserved site
			GRMZM2G117335	138676794	138679563	-1	IPR000104	Antifreeze protein, type I
6	**138743574**	**82.73**	**GRMZM2G085964**	**138736793**	**138738075**	**-1**	**IPR001471| IPR016177|**	**Pathogenesis-related transcriptiol factor and ERF, D-binding| D-binding, integrase-type|**

Table 2.3 (Continued)

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
6	139162942	82.49| 82.49	GRMZM2G049342	139105345	139120312	-1	IPR000403| IPR003151| IPR009076| IPR011009| IPR016024	Phosphatidylinositol 3- and 4-kise, catalytic| PIK-related kise, FAT| FKBP12-rapamycin-associated protein, FKBP12-rapamycin-binding| Protein kise-like| Armadillo-type fold
6	148277163| 148277351| 148277188| 148277458	84.7| 83.69| 83.69| 82.2	GRMZM2G032865	148275618	148277630	1	na	unknown
6	148261988	82.2	GRMZM2G058472	148135300	148140036	-1	IPR002495	Glycosyl transferase, family 8
7	117921262	12.6	GRMZM2G096092	117913011	117916163	1	na	unknown
7	101332276	11.06	GRMZM2G450488	101315161	101321984	1	IPR003663| IPR005828| IPR005829| IPR011701|	Sugar/inositol transporter| General substrate transporter| Sugar transporter, conserved site| Major facilitator superfamily MFS-1|
7	99932263	10.71	GRMZM2G476685	99929088	99933905	-1	IPR000104| IPR004367| IPR006670| IPR006671| IPR011028	Antifreeze protein, type I| Cyclin, C-termil| Cyclin| Cyclin, N-termil| Cyclin-like
7	92027603	10.25	GRMZM2G041631	91372693	91374911	1	IPR000109| IPR016196|	TGF-beta receptor, type I/II extracellular region| Major facilitator superfamily, general substrate transporter|
7	122145665	10.21| 10.21	GRMZM2G140633	122142128	122144701	1	IPR004367| IPR006670| IPR006671| IPR011028|	Cyclin, C-termil| Cyclin| Cyclin, N-termil| Cyclin-like|
7	138399286	18.21	GRMZM2G052288	138293515	138295717	-1	IPR000183| IPR002433|	Orn/DAP/Arg decarboxylase 2| Ornithine decarboxylase|
7	136306019	18.13	GRMZM2G354525	136286786	136298679	-1	IPR005045| IPR012336|	unknown function DUF284, transmembrane eukaryotic| Thioredoxin-like fold|
			GRMZM2G052610	136305883	136310364	1	IPR001763	Rhodanese-like
			GRMZM2G052509	136309715	136312366	-1	IPR001806| IPR002041| IPR003579| IPR005225| IPR006688| IPR006689| IPR013684| IPR013753|	Ras GTPase| Ran GTPase| Ras small GTPase, Rab type| Small GTP-binding protein| ADP-ribosylation factor| ARF/SAR superfamily| Miro-like| Ras|
7	137768235	17.49	GRMZM2G109560	137749220	137758057	-1	IPR001270| IPR003593| IPR003959| IPR008268| IPR011703| IPR016137|	Chaperonin clpA/B| ATPase, AAA+ type, core| ATPase, AAA-type, core| Peptidase S16, active site| ATPase associated with various cellular activities, AAA-3| Regulator of G protein sigling superfamily|
7	137238098	17.16	GRMZM2G074472	137235668	137237420	1	na	unknown
7	137209509	17.13| 17.13	GRMZM2G048131	137198544	137200001	1	IPR003311	AUX/IAA protein
			GRMZM2G048170	137206975	137208187	-1	na	unknown
8	141758758| 141768482| 141767370	125.7| 124.6| 124.6	GRMZM2G466848	141758273	141758762	1	na	unknown
8	141790324	124.7	GRMZM2G005304	141769237	141769860	-1	na	unknown
8	141769588| 141768482| 141767370	124.6| 124.6| 124.6	GRMZM2G005304	141769237	141769860	-1	na	unknown
8	149500562| 149498905| 149500785	124.0| 122.6| 122.3	GRMZM2G009732	149495340	149495988	-1	na	unknown
			GRMZM2G311680	149497572	149499826	-1	IPR001461| IPR001859| IPR001969	Peptidase A1| Ribosomal protein P2| Peptidase aspartic, active site
8	149528832| 149536625	122.5| 122.5	GRMZM2G462623	149528498	149534012	1	IPR014889	Transcription factor DP, C-termil
		122.5| 122.5	GRMZM2G163988	149534996	149537085	-1	IPR002583	Ribosomal protein S20
9	**12502229**	**4.951**	**GRMZM2G428370**	**12481725**	**12484816**	**1**	**IPR001245| IPR001611| IPR002290| IPR017441|**	**Tyrosine protein kise| Leucine-rich repeat| Serine/threonine protein kise| Protein kise, ATP binding site|**
			GRMZM2G127990	**12502375**	**12505660**	**1**	**IPR001245| IPR001611| IPR002290| IPR017441|**	**Tyrosine protein kise| Leucine-rich repeat| Serine/threonine protein kise| Protein kise, ATP binding site|**
9	**11439100**	**4.801**	**GRMZM2G171179**	**11409298**	**11412258**	**-1**	**IPR001471| IPR016177|**	**Pathogenesis-related transcriptiol factor and ERF, D-binding| D-binding, integrase-type|**
			GRMZM2G024530	11439405	11444381	1	IPR001092| IPR011598|	Basic helix-loop-helix dimerisation region bHLH| Helix-loop-helix D-binding|
9	12795420| 12798504	3.851| 3.849	GRMZM2G380205	12725157	12731505	1	IPR003593| IPR017871|	ATPase, AAA+ type, core| ABC transporter, conserved site|
			GRMZM2G067053	12799430	12804267	1	IPR000210| IPR004249| IPR011333	BTB/POZ-like| NPH3| BTB/POZ fold
9	11509292	3.818| 3.818	GRMZM2G046618	11506823	11509549	-1	IPR006501| IPR011050|	Pectinesterase inhibitor| Pectin lyase fold/virulence factor|
10	72405736	12.37	GRMZM2G141071	72381934	72382743	-1	na	unknown

Table 2.3 (Continued)

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
10	70808287	12.36	GRMZM2G112782	70795096	70798921	-1	IPR010989| IPR015260|	t-SRE| Syntaxin 6, N-termil|
10	35606083	11.6	GRMZM2G009443	35586545	35593339	1	IPR001270| IPR001943| IPR003593| IPR003959| IPR004176| IPR013093|	Chaperonin clpA/B| UvrB/UvrC protein| ATPase, AAA+ type, core| ATPase, AAA-type, core| Clp, N-termil| ATPase associated with various cellular activities, AAA-2|
10	26764069	11.54	GRMZM2G078508	26760967	26763809	-1	IPR000526| IPR011051| IPR013096	Auxin-binding protein| Cupin, RmlC-type| Cupin 2, conserved barrel
10	26723785	11.54| 11.54	GRMZM2G316665	26718334	26721032	-1	na	unknown
		11.54| 11.54	GRMZM2G316593	26726418	26727788	-1	IPR009053	Prefoldin
10	130536818	10.63	GRMZM2G446557	130532827	130536392	1	IPR005269	Conserved hypothetical protein CHP00730
			GRMZM2G063380	130543295	130546305	-1	IPR002885	Pentatricopeptide repeat
10	136915251| 136915096	10.06| 9.055	GRMZM2G094860	136913045	136915128	1	IPR006214	Uncharacterised protein family UPF0005
10	135748030| 135748031	9.35| 9.128	GRMZM2G040673	135744333	135746056	-1	IPR001904| IPR003072| IPR003980| IPR006458|	Paxillin| Orphan nuclear receptor, NOR1 type| Histamine H3 receptor| unknown function DUF623, plant|
			GRMZM2G040359	135746269	135748634	-1	IPR001087| IPR013830|	Lipase, GDSL| Esterase, SGNH hydrolase-type|

Other intervals of interest:

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
3	160697657| 160697324| 160695996	30.72| 30.72| 28.19	GRMZM2G153991	160691195	160693947	1	IPR001179	Peptidyl-prolyl cis-trans isomerase, FKBP-type
3	160805662| 160805601	27.94| 27.94	GRMZM2G180406	160801253	160804532	1	IPR001092| IPR011598|	Basic helix-loop-helix dimerisation region bHLH| Helix-loop-helix D-binding|
3	212040830	9.799	GRMZM2G702221	212032991	212033206	1	na	unknown
			GRMZM2G004183	212039912	212042050	1	IPR004253	unknown function DUF231, plant
3	213901128| 213896123	9.492| 8.83	GRMZM2G121878	213888899	213896251	-1	IPR001765	Carbonic anhydrase
			GRMZM2G122028	213900646	213902423	-1	IPR012442	unknown function DUF1645
3	213890434| 213896123	9.487| 8.83	GRMZM2G121878	213888899	213896251	-1	IPR001765	Carbonic anhydrase
3	212254248	8.618| 8.618	GRMZM2G349407	212244596	212246542	-1	na	unknown
			GRMZM2G047129	212247706	212251776	-1	IPR001087| IPR013830|	Lipase, GDSL| Esterase, SGNH hydrolase-type|
4	181142581	29.82	GRMZM2G025806	181142535	181146752	-1	IPR005176| IPR009060|	unknown function DUF298| UBA-like|
4	180345438| 180343020	28.18| 26.10	GRMZM2G016948	180344442	180346058	-1	IPR007650	unknown function DUF581
4	180425560	25.76	GRMZM2G016948	180344442	180346058	-1	IPR007650	unknown function DUF581
4	185918871	23.3| 23.3	GRMZM2G142777	185913119	185915014	-1	IPR013078	Phosphoglycerate mutase
			AC204868.3_FG004	185917632	185918929	1	IPR001810	Cyclin-like F-box
			GRMZM2G076087	185925082	185927639	-1	IPR003256| IPR005824| IPR008991	Ribosomal protein L24| KOW| Translation protein SH3-like
5	189569948	18.2	GRMZM2G169458	189566293	189570723	-1	IPR015590| IPR016161|	Aldehyde dehydrogese| Aldehyde/histidinol dehydrogese|
			GRMZM2G169558	189574407	189575999	-1	IPR001245| IPR002290| IPR011009	Tyrosine protein kise| Serine/threonine protein kise| Protein kise-like
5	**190880589| 190871333**	**17.50| 17.04**	**GRMZM2G057386**	**190879782**	**190881307**	**1**	**IPR001471| IPR016177|**	**Pathogenesis-related transcriptiol factor and ERF, D-binding| D-binding, integrase-type|**
			GRMZM2G057231	190887149	190890489	-1	IPR015928	Aconitase/3-isopropylmalate dehydratase, swivel
5	190422625	16.5	GRMZM2G124136	190401253	190404005	1	IPR011701	Major facilitator superfamily MFS-1
5	191109424	16.39| 16.39	GRMZM2G139920	191072558	191076067	1	IPR004841| IPR018130|	Amino acid permease-associated region| Ribosomal protein S2, conserved site|
			GRMZM2G056903	191110266	191112774	-1	IPR001245| IPR002290| IPR011009	Tyrosine protein kise| Serine/threonine protein kise| Protein kise-like
			GRMZM2G056889	191115112	191116577	-1	na	unknown
5	**203735206| 203735229**	**16.88| 16.39**	**GRMZM2G019317**	**203723478**	**203727851**	**-1**	**IPR001245| IPR001611| IPR002290| IPR017441|**	**Tyrosine protein kise| Leucine-rich repeat| Serine/threonine protein kise| Protein kise, ATP binding site|**
5	201961545| 201961623| 201961544	16.02| 15.22| 15.22	GRMZM2G180930	201948858	201952213	1	IPR002961| IPR005946|	Tumour necrosis factor c/lymphotoxin-beta| Phosphoribosyl pyrophosphokise|
			GRMZM2G180926	201954299	201955863	1	IPR000169| IPR000668| IPR013201	Peptidase, cysteine peptidase active site| Peptidase C1A, papain C-termil| Proteise inhibitor I29, cathepsin propeptide

Table 2.3 (Continued)

chr	SNP(s) (AGP 4a.53)	log(1/p)	maize gene	start	end	strand	interpro domain id(s)	domain name(s)
			GRMZM2G180922	201956451	201961027	1	IPR001509| IPR002198| IPR002225| IPR003869| IPR005913| IPR013120| IPR016040	D-dependent epimerase/dehydratase| Short-chain dehydrogese/reductase SDR| 3-beta hydroxysteroid dehydrogese/isomerase| Polysaccharide biosynthesis protein CapD| dTDP-4-dehydrorhamnose reductase| Male sterility, D-binding| D(P)-binding
			GRMZM2G180920	201962624	201964070	1	na	unknown
7	140184553	19.67	GRMZM2G375307	140161276	140165654	-1	IPR001005| IPR006578|	SANT, D-binding| MADF domain|
			GRMZM2G006416	140187830	140193578	1	IPR000222	Protein phosphatase 2C, manganese/magnesium aspartate binding site
7	146802842| 146802876| 146802990	19.48| 19.48| 19.48	GRMZM2G076225	146793319	146797012	1	IPR001117| IPR008972|	Multicopper oxidase, type 1| Cupredoxin|
			GRMZM2G076387	146803324	146807841	1	na	unknown
7	142893587	19.47| 19.47	GRMZM2G111309	142888881	142891479	-1	IPR000120	Amidase sigture enzyme
			GRMZM2G413044	142899087	142900845	1	IPR000757| IPR008985| IPR016455	Glycoside hydrolase, family 16| Concavalin A-like lectin/glucase| Xyloglucan endotransglucosylase/hydrolase
8	147688291	132.6	GRMZM2G007848	147609001	147615308	-1	IPR001245| IPR002290| IPR003527| IPR008350| IPR017441	Tyrosine protein kise| Serine/threonine protein kise| MAP kise, conserved site| ERK3/4 MAP kise| Protein kise, ATP binding site
8	144999684	129.5	GRMZM2G034025	144989059	145003544	1	IPR006155| IPR011046|	Machado-Joseph disease protein MJD| WD40 repeat-like|
8	148788680	129.4	GRMZM2G163159	148766012	148779545	-1	na	unknown
			GRMZM2G163157	148782639	148784127	-1	IPR001005| IPR006578|	SANT, D-binding| MADF domain|
8	146930527	128.3	GRMZM2G030167	146872954	146882246	1	IPR002842	ATPase, V1/A1 complex, subunit E
8	148221345	128.1| 128.1	GRMZM2G420935	148034068	148034643	1	IPR003579	Ras small GTPase, Rab type
9	143532815| 143532785| 143533469| 143533378	5.863| 5.863| 5.507| 5.507	GRMZM2G109997	143526087	143532813	1	IPR002891	Adenylylsulphate kise, C-termil
			GRMZM2G048626	143537296	143539413	1	IPR017853	Glycoside hydrolase, catalytic core
9	143201854	5.381| 5.381	GRMZM2G102802	143198247	143208265	1	na	unknown
10	114570088| 114570157| 114570261	12.20| 12.20| 12.20	GRMZM2G027241	114494684	114497241	1	IPR004345	TB2/DP1 and HVA22 related protein
10	114611113| 114611321	12.2| 12.2	GRMZM2G027241	114494684	114497241	1	IPR004345	TB2/DP1 and HVA22 related protein

Three of the NAM QTL had SNP associations with genes containing predicted kinase - LRR domains. A fourth kinase-LRR - SNP association was located just outside of the QTL confidence interval on chromosome 5. While genes with NB-LRR domains have long been implicated in plant disease resistance and could be considered canonical 'R-genes' (Sacco and Moffett 2009), the involvement of LRR related genes in quantitative resistance has been the subject of much debate (Poland *et al.* 2009). The rice resistance genes *Xa21* and *Xa3*/*Xa6* contain kinase and LRR domains (Andaya and Ronald 2003; Sun *et al.* 2004). These genes, known as receptor-like kinases (RLKs), differ from typical R-genes in that they lack a distinctive NB domain. Recently, research in plant basal immunity has implicated several RLKs in plant pathogen defense. *FLS2* and *ERF1* are involved in recognition of bacterial flagellin and Ef-Tu, respectively, and initiate basal defense (Gomez-Gomez and Boller 2000). It has been hypothesized that quantitative resistance could share commonalities with basal defense (Poland *et al.* 2009) and that the non-canonical RLK R-genes could primarily function as pathogen-associated molecular pattern receptors (Sacco and Moffett 2009). The implication of four different RLKs in quantitative resistance suggests that this class of genes may be broadly associated with the loci of small effect that condition QDR.

At two QTL, SNP association fell in or upstream of pathogen defense related ethylene response factor (*ERF*). *ERF1* has been broadly implicated in resistance to necrotrophic pathogens (Berrocal-Lobo *et al.* 2002). Activated at the convergence of the ethylene and jasmonate pathways, *ERF1* is an integral component of defense activation (Lorenzo *et al.* 2003). The identification of

SNPs that localize to the upstream regions of these two genes suggests a role for polymorphisms that affect transcriptional regulation.

Association with an *Mlo*-like gene was found for a QTL on the telomeric region of Chr. 2. *Mlo* genes are known to be membrane associated and consist of seven trans-membrane domains. A second gene with multiple trans-membrane domains was within 5kb of the *Mlo*-like gene. *Mlo* was first identified in barley. The recessive *mlo* mutation confers broad-spectrum resistance to *Erysiphe graminis* f. sp. *hordei*, the causal pathogen of powdery mildew disease (Büschges *et al.* 1997). It is thought that functional *Mlo* protein is needed as a compatibility factor for the fungus (Matt *et al.* 2006).

A SNP in an ABC-transporter like gene was associated with resistance on Chr.7. The recently identified wheat rust gene, *Lr34*, is a putative ABC-transporter (Krattinger *et al.* 2009). The clear connection of other candidate genes to disease resistance from this association study is limited by a lack of knowledge of the molecular basis for quantitative disease resistance. There are also apparent limitations for using NAM in gene finding. This is apparent on Chr. 8 where two closely liked QTL of large effects confound SNP association. The profile of SNP associations on Chr.8 is clearly driven by linkage. We are also working with a limited number of SNPs for genome-wide coverage in maize. An order of magnitude higher SNP coverage (~10M) will be needed to sufficiently capture LD through the entire maize genome.

Quantitative disease resistance remains an important research frontier. By leveraging publicly available community resources, we were able to map over 220 resistance alleles at 30 QTL for an economically important disease in maize. Using nested association mapping we were able to link six defense related genes to QDR, increasing the understanding of this trait. We

are undertaking further studies using near isogenic lines to confirm the QTL effects. Association mapping in a larger, more diverse inbred line panel is also underway to confirm the trait-gene associations.

Despite the importance of QDR in crop production and ecology, the genetic basis of QDR remains largely unknown. With the genetic resources available in maize, we see this important crop species emerging as a model system for the study of QDR. Coupled with a strong history of robust large-scale field trials and community expertise in quantitative genetics, maize pathosystems can serve as an excellent tool to unravel the genetics underlying QDR. Uncovering the molecular mechanisms of complex disease resistance in plants will assist in the development of durably resistant crop cultivars, increasing food security.

MATERIALS AND METHODS

Plant Material: The maize nested association mapping (NAM) population consists of 25 recombinant inbred line (RIL) families derived from crossing each of 25 diverse founder inbred lines to the reference parent B73 followed by 5 generations of self mating (Yu *et al.* 2008; Buckler *et al.* 2009; McMullen *et al.* 2009). The public intermated B73 x Mo17 (IBM) population was also included as a 26th family. Each family consists of 200 S_5 RILs, of which 4,631 were used for this study, averaging 178 lines per cross. The RILs are genotyped with 1106 SNP markers (McMullen *et al.* 2009).

Phenotypic evaluation for resistance to NLB: Field trials were conducted during 2007, 2008, and 2009 at the Robert B. Musgrave Research Farm in Aurora, New York. Trials were planted on May 15, 2007; May 14, 2008 and May 18, 2009. Lines were planted as single row plots 7` in length

with 30” between rows. Plots were overplanted and thinned to 10 plants/row.

Trials were laid out in an augmented incomplete block design with one replication in each year. For each trial, lines were grouped by family with augmented incomplete blocks within each family. Each incomplete block consisted of 20 RILs and two checks: B73 and the second parent for the respective family.

All trials were artificially inoculated with *E. turcicum* Race 1 (isolate NY-001 from the lab of R. Nelson). Plants were inoculated at about the 6-8 leaf stage which corresponded to July 2, 2007; June 27, 2008 and June 16, 2009. Each year, two types of inoculums were simultaneously applied: 2.5–3.0 ml of dried infected sorghum grains, previously inoculated and cultured for 2 weeks, and 1.0 ml of a spore suspension (1×10^3 conidia/ml) in H_2O with 0.02% Tween 20. For the spore suspension, *E. turcicum* isolate NY001 was cultured on lactose casein agar (LCA) plates for 2-3 weeks at room temperature under 12 hr light / 12 hr dark condition. Spores were harvested by flooding the plates with sterile H_20 and scraping with a glass rod and filtering through cheese cloth. Spore concentrations were determined and diluted to 1×10^3 conidia/ml.

In 2007, two disease phenotypes, incubation period (IP) and disease severity (DS), were evaluated, while in 2008 and 2009 only DS was evaluated due to the extreme time requirements for evaluation of IP. For the evaluation of IP, all plots were evaluated daily. IP was measured as the number of days after inoculation that the first water-soaked lesion was observed on 50% of the plants in a row. For DS, all plots were evaluated at three time points during the season at 10 day intervals, with the first rating corresponding to shortly after anthesis for B73. DS ratings were conducted by visually evaluating each

plot and rating the percentage of total diseased leaf area (DLA) using a 0-100% rating scale with 1% increments.

Additional phenotypes collected included days to anthesis (DTA) measured as the day when 50% of the plants in a row were shedding pollen and plant and ear height measured in cm to the base of the flag leaf and base of the primary ear bearing node. In 2007, high levels of natural smut infection (*Ustilago maydis*) were observed and recorded as the number of infected plants in a row (incidence) and the amount of smut on those plants (severity).

Statistical Analysis: The trait distribution for DLA was skewed toward resistance, so a square root transformation was employed to normalize the trait distribution before further analysis was conducted. To account for year and field effects a multivariate mixed model was run in ASReml (VSN International). The model is as follows:

$$sqrtDLA_{hijklm} = T_h + TY_{hi} + TP_{hj} + TYP_{hij} + TpB_{hk(i)} + TB_{hl(k)} + TL_{hm(j)} + TYPL_{hi(j)m}$$

where sqrtDLA is the square root transformed DLA rating for a recombinant inbred line *m* in population *j*, at time *h*, in year *i*, within population block *k*, and incomplete block *l*. T_h is a fixed effect of rating *h*. TY_{hi} is the fixed effect of year *i* at rating *h*. TP_{hi} is the random effect of population *j* at rating *h*. TYP_{hij} is the random interaction of year and population at rating *h*. $TpB_{hk\,(i)}$ is the effect of population block within year at a given rating. $TB_{hl(k)}$ is the random effect of incomplete block within population block at rating *h*. $TL_{hm(j)}$ is the random effect of RIL *m* in population *j* at rating *h*, and $TYPL_{hi(j)m}$ is the random interaction of year, population and line at a given rating. A unique covariance

structure was specified for each combination of rating time with year, population, and population block within year. RIL-by-rating variances were also modeled uniquely for each population. All founder inbred lines were considered a separate population and modeled accordingly. Model solutions provided best linear unbiased predictors (BLUPs) for each RIL at each of the three ratings. An NLB index was calculated by averaging the three disease ratings.

Broad sense heritabilities for the line BLUPs were estimated for the whole NAM and for each individual family. With the multivariate response, heritability was calculated for each rating and for the combined NLB index. Heritability on a line mean basis was calculated as:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{ge}^2}{e} + \frac{\sigma_e^2}{r}}$$

where σ_g^2 is the family genetic variance, σ_{ge}^2 is the genotype x environment, σ_e^2 is the average residual error variance, *e* is the harmonic mean number of environments lines were evaluated in, and *r* is harmonic mean number of replications of a given line within environment. The same formula was used to calculate H^2 in NAM with the following modification: σ_g^2 is the sum of between family variance and the average family genetic variance. H^2 for the NLB index was calculated according to Lin and Allaire (Lin and Allaire 1977) as:

$$H^2 = \frac{\mathrm{b'Gb}}{\mathrm{b'Pb}}$$

where b is the vector of index coefficients, G is the genetic variance-covariance matrix and P is the phenotypic variance-covariance matrix.

Joint general linear model: To identify markers associated with QTL, a joint general linear model was selected using stepwise model selection with Proc GLMselect in SAS v9.1.3 software. Each individual family was fit as a

main effect. Markers were coded as 0, 1 and 2 for B73, heterozygous and second parent alleles, respectively, according to an additive genetic model. Markers were nested within population and stepwise selection was conducted with effect selection/removal set at 1×10^{-4} as determined by permutation analysis described below. To avoid selection of maturity related effects, days to anthesis was included as a covariate in model selection (Buckler *et al.* 2009). Selected marker effects were then fit into a GLM and dropped individually to confirm significance. The joint GLM has limited power to detect QTL represented in a single population. Therefore a final selection step was conducted. The full model including population and selected marker effects was fit. Stepwise GLM was then conducted on an individual population basis with the same selection threshold of the marginal F-test p-value $< 10^{-4}$.

To correct for multiple testing, permutation analysis was conducted by randomizing the NLB index values within each population and then identifying the most significant marker effect using stepwise selection as described above. This was repeated for 1000 iterations to determine a selection threshold of p-value = 10^{-4} corresponded to an experimental alpha=0.05.

To best model the data from evaluation of IP, a Cox proportional hazards model was fit in R statistical software (R.Development.Core.Team 2009). Stepwise model selection was conducted with selection of marker effects determined by the p-value from a Wilk's F-test. The threshold for effect selection was $p < 10^{-4}$.

For each QTL position identified, confidence intervals were constructed by sequentially examining flanking markers. Starting with the first flanking marker, the QTL marker and the flanking marker were both fit into the full linear model. The p-value from the marginal F-test for the QTL marker was

then determined. If the QTL marker did not have a significant contribution to the model (p-value<0.05), the flanking marker was considered equivalent to the QTL marker and within the 95% confidence interval. The flanking marker was then moved outward and the test repeated until the QTL marker significantly contributed to the model. To compensate for regions of low marker density, pseudo-markers were imputed where consecutive markers were farther apart than 1cM. Pseudo-markers were imputed at 1cM intervals between flanking markers based on the expected marker class at that position.

The same model selection procedure was conducted for a multivariate model with the trait values from the three ratings as the response variables. Model selection was programmed in R statistical software (R.Development.Core.Team 2009) with effect selection determined by Wilk's F-test. Permutation analysis was conducted as described above for 200 permutations. The selection threshold of 10^{-7} was determined to correspond to an experimental alpha of 0.05.

Genomewide SNP Association: Single nucleotide polymorphisms (SNPs) from the maize hap-map v.1 (http://www.maizegenetics.net/maize-hap-map) for the 25 founder lines, Mo17 and B73 were tested for association with NLB resistance. Missing SNP genotypes were imputed using fastPHASE software (F. Tian, P. Bradbury et.al unpublished) (Scheet and Stephens 2006). Marker positions were referenced to the B73 AGP v1 physical map. SNP markers were coded as 0 and 1 for B73 and non-B73 polymorphisms, respectively. The complete SNP dataset was then imputed to the NAM RILs. SNPs that were polymorphic between B73 and a given founder inbred were imputed to the RILs of that respective family. For a given SNP, the imputed value was based on the expected genotype at the physical location of the SNP

based on the relative distance from the physical position of the closest NAM flanking markers. SNPs that were not polymorphic between B73 and a given founder were imputed as 0 to all progeny of that family.

Trait-marker association was conducted by chromosome. Residual phenotypic values for each chromosome were determined by fitting the full marker model and then sequentially dropping the markers from each chromosome. The residual values from this reduced model were assigned to the respective chromosome and used as the phenotypic trait value from marker associations. Trait-marker association for each SNP was evaluated for each chromosome by fitting a linear model between the imputed SNP genotypes and the corresponding residual phenotypes. The strength of trait-marker association was determined from the percentile of a respective F-distribution. From the F-test, log (1/p-value) was plotted against the physical position of the SNP. Within QTL confidence intervals, the five SNP markers with the highest log (1/p) values were examined as polymorphisms in LD with potential candidate genes from the B73 filtered gene set (MaizeSequence Release 4a.53 - www.maizesequence.org).

APPENDIX 1

Association between relative maturity and disease resistance.

For several necrotrophic maize diseases including NLB, a negative association between relative maturity and disease resistance is consistently observed. This observation is generally reported for trials where disease and maturity are recorded on the same plots leading to questions of causality. Does early maturity lead to increased disease, or do high disease levels lead to plants under stress and early flowering? To examine this in NAM we compared data from 2007 and 2008 where an un-inoculated NAM trial was planted on the same research farm on the same date as the inoculated trial. The flowering time difference between inoculated and un-inoculated trials was examined by comparing the difference in flowering date between the two trials and fitting the difference against disease severity. The was no observable trend in this test, indicating that the association between relative maturity and disease severity is largely due to maturity effects on resistance, rather than the other way around.

REFERENCES

Andaya, C. B. and P. C. Ronald (2003). A Catalytically Impaired Mutant of the Rice *Xa21* Receptor Kinase Confers Partial Resistance to *Xanthomonas oryzae* Pv *Oryzae*. *Physiological and Molecular Plant Pathology* **62** (4): 203-208.

Ayliffe, M., R. Singh and E. Lagudah (2008). Durable Resistance to Wheat Stem Rust Needed. *Current Opinion in Plant Biology* **11** (2): 187-192.

Berrocal-Lobo, M., A. Molina and R. Solano (2002). Constitutive Expression of *Ethylene-Response-Factor1* in Arabidopsis Confers Resistance to Several Necrotrophic Fungi. *The Plant Journal* **29**: 23-32.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J. C. Glaubitz, M. M. Goodman, C. Harjes, K. Guill, D. E. Kroon, S. Larsson, N. K. Lepak, H. Li, S. E. Mitchell, G. Pressoir, J. A. Peiffer, M. O. Rosas, T. R. Rocheford, M. C. Romay, S. Romero, S. Salvo, H. S. Villeda, H. Sofia da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich and M. D. McMullen (2009). The Genetic Architecture of Maize Flowering Time. *Science* **325** (5941): 714-718.

Büschges, R., K. Hollricher, R. Panstruga, G. Simons, M. Wolter, A. Frijters, R. van Daelen, T. van der Lee, P. Diergaarde, J. Groenendijk, S. Töpsch, P. Vos, F. Salamini and P. Schulze-Lefert (1997). The Barley *Mlo* Gene: A Novel Control Element of Plant Pathogen Resistance. *Cell* **88** (5): 695-705.

Flint-Garcia, S. A., A.-C. Thuillet, J. Yu, G. Pressoir, S. M. Romero, S. E. Mitchell, J. Doebley, S. Kresovich, M. M. Goodman and E. S. Buckler

(2005). Maize Association Population: A High-Resolution Platform for Quantitative Trait Locus Dissection. *Plant Journal* **44** (6): 1054-1064.

Fu, D., C. Uauy, A. Distelfeld, A. Blechl, L. Epstein, X. Chen, H. Sela, T. Fahima and J. Dubcovsky (2009). A Kinase-Start Gene Confers Temperature-Dependent Resistance to Wheat Stripe Rust. *Science* **323** (5919): 1357-1360.

Fukuoka, S., N. Saka, H. Koga, K. Ono, T. Shimizu, K. Ebana, N. Hayashi, A. Takahashi, H. Hirochika, K. Okuno and M. Yano (2009). Loss of Function of a Proline-Containing Protein Confers Durable Disease Resistance in Rice. *Science* **325** (5943): 998-1001.

Gomez-Gomez, L. and T. Boller (2000). *FLS2*: An LRR Receptor-Like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. *Molecular Cell* **5** (6): 1003-1011.

Krattinger, S. G., E. S. Lagudah, W. Spielmeyer, R. P. Singh, J. Huerta-Espino, H. McFadden, E. Bossolini, L. L. Selter and B. Keller (2009). A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. *Science* **323** (5919): 1360-1363.

Lee, M., N. Sharopova, W. D. Beavis, D. Grant, M. Katt, D. Blair and A. Hallauer (2002). Expanding the Genetic Map of Maize with the Intermated B73 x Mo17 (Ibm) Population. *Plant Molecular Biology* **48** (5-6): 453-461.

Lin, C. Y. and F. R. Allaire (1977). Heritability of a Linear Combination of Traits. *TAG Theoretical and Applied Genetics* **51** (1): 1-3.

Lorenzo, O., R. Piqueras, S. J. J. Sanchez and R. Solano (2003). *Ethylene Response Factor1* Integrates Signals from Ethylene and Jasmonate Pathways in Plant Defense. *Plant Cell* **15** (1): 165-178.

Matt, H., C. Chiara and P. Ralph (2006). *Mlo*-Based Powdery Mildew Immunity: Silver Bullet or Simply Non-Host Resistance? *Molecular Plant Pathology* **7** (6): 605-610.

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S. E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M. O. Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J. C. Glaubitz, M. Goodman, D. Ware, J. B. Holland and E. S. Buckler (2009). Genetic Properties of the Maize Nested Association Mapping Population. *Science* **325** (5941): 737-740.

Poland, J. A., P. J. Balint-Kurti, R. J. Wisser, R. C. Pratt and R. J. Nelson (2009). Shades of Gray: The World of Quantitative Disease Resistance. *Trends in Plant Science* **14** (1): 21-29.

R Development Core Team (2009). R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Sacco, M. A. and P. Moffett (2009). Disease Resistance Genes: Form and Function. Molecular Plant-Microbe Interactions. K. Bouarab, N. Brisson and F. Daayf. Wallingford, UK, , CABI**:** 94-141.

Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu, K. A. Fengler, R. Meeley, E. V. Ananiev, S. Svitashev, E. Bruggemann, B. Li, C. F. Hainey, S. Radovic, G. Zaina, J. A. Rafalski, S. V. Tingey, G.-H. Miao, R. L. Phillips and R. Tuberosa (2007). Conserved Noncoding Genomic Sequences Associated with a Flowering-Time Quantitative Trait Locus in Maize. *Proceedings of the National Academy of Sciences* **104** (27): 11376-11381.

Scheet, P. and M. Stephens (2006). A Fast and Flexible Statistical Model for Large-Scale Population Genotype Data: Applications to Inferring Missing Genotypes and Haplotypic Phase. **78** (4): 629-644.

Sun, X., Y. Cao, Z. Yang, C. Xu, X. Li, S. Wang and Q. Zhang (2004). *Xa26*, a Gene Conferring Resistance to *Xanthomonas oryzae* Pv. *Oryzae* in Rice, Encodes an LRR Receptor Kinase-Like Protein. *The Plant Journal* **37**: 517-527.

Wisser, R. J., P. J. Balint-Kurti and R. J. Nelson (2006). The Genetic Architecture of Disease Resistance in Maize: A Synthesis of Published Studies. *Phytopathology* **96** (2): 120-129.

Yu, J., J. B. Holland, M. D. McMullen and E. S. Buckler (2008). Genetic Design and Statistical Power of Nested Association Mapping in Maize. *Genetics* **178** (1): 539-551.

CHAPTER 3

Dissection of multiple disease resistance in maize using nested association mapping

ABSTRACT

In natural systems and in crop production environments, plants are challenged with numerous pathogen species and races. To be successful, an individual plant or a cultivar must be able to defend itself against a diverse array of attackers. Through artificial selection, plant cultivars have been developed that display multiple disease resistance (MDR). While a few genes that condition resistance to multiple pathogens have recently been identified, the genetic architecture of MDR remains poorly understood. Insights into the genetic architecture that underlies the MDR phenotype will be of value to plant breeding and ecology. Here we present a genetic analysis of the MDR phenotype based on the maize Nested Association Mapping (NAM) population, a set of 5,000 recombinant inbred lines derived from 26 diverse inbreds. The NAM population was evaluated for resistance to three economically important maize diseases: southern leaf blight, northern leaf blight and gray leaf spot. The diverse founder inbred lines showed a strong correlation for resistance to the different diseases. The recombinant inbred progeny also had correlated resistance, though the effect was weaker. Within individual families there were significant, though weak, phenotypic correlations between diseases. This progressive decrease in correlated resistance indicated that most of the MDR phenotype was due to combining many disease specific resistance genes, but some evidence was obtained for pleiotropy or linkage of genes affecting more than one disease. Twenty-three

genetic positions were identified where quantitative resistance loci for two or more diseases co-localized. To examine the possibility of MDR genes, the estimated allele effects from each founder inbred were compared. At seven loci, positively correlated allele effects provided evidence for MDR genes. Using standardized trait values, an MDR index was calculated and used for mapping potential MDR QTL. The MDR index was then used to test >1.6M SNPs for association with multiple disease resistance. Three candidate genes with known roles in disease resistance were at or in proximity to associated SNPs. Further characterization of these regions will be needed to confirm that these genes condition resistance. Analysis of the NAM population suggested that resistance to the three diseases studied here is largely due to the accumulation of disease-specific genes and, to a limited extent, pleiotropic genes that condition MDR.

INTRODUCTION

Plants are challenged by a range of microbes that can cause disease. These diverse pathogens have adopted various mechanisms to reproduce on their plant hosts. In turn, plants have adopted an array of defense mechanisms to mitigate infection and disease development. It is expected that any genetic mechanism that confers resistance to multiple pathogens would contribute to fitness for a plant species and, consequently, be a strong target for natural and artificial selection. Genes conditioning natural variation in multiple disease resistance (MDR) would have considerable importance for crop improvement.

In contrast to qualitative resistance genes, which typically confer resistance to one or a few races of a single pathogen, quantitative resistance genes (QDR) genes generally confer race-non-specific resistance (Poland *et al.* 2009). It has been hypothesized that some QDR genes could be involved in more general forms of resistance and confer resistance to multiple different pathogens (Poland *et al.* 2009). Indeed, genes with pleiotropic effects on resistance to multiple pathogens have been cloned. The QDR gene, *Lr34*, a putative ATP-binding cassette transporter protein, confers resistance to leaf rust, stripe rust, and powdery mildew. The MDR mechanism for *Lr34* is yet to be determined (Krattinger *et al.* 2009). Additional genes identified in Arabidopsis have been implicated in MDR through mutant studies (Century *et al.* 1995; Tierens *et al.* 2002; Mou *et al.* 2003; Murray *et al.* 2005). Several loci have been found to contribute resistance to a category of pathogen (e.g. those with a necrotrophic lifestyle) while conditioning increased susceptibility to a contrasting group (e.g. those with a biotrophic lifestyle) (e.g.Lorang *et al.* 2007; Nagy *et al.* 2007).

In the context of plant breeding, MDR has been a focus of selection in programs where cultivars target regions where environmental conditions favor disease. This has led to the creation of cultivars that exhibit an MDR phenotype. This phenotype contributes to yield stability in the presence of high disease pressure for a range of diseases. There are several scenarios which could describe the underlying architecture of this MDR phenotype. The first scenario would result from the accumulation of disease-specific resistance genes. This scenario can be further specified to a second scenario which would incorporate clusters of disease specific genes. The final scenario would entail the selection of genes of pleiotropic effects. These scenarios are not mutually exclusive as a combination of any or all of these architectures could lead to the resistance phenotype. These scenarios are difficult to distinguish as most mapping studies have limited resolution to distinguish linkage from pleiotropy.

Correlated levels of resistance to different diseases have been noted in germplasm collections, leading some authors to suggest the possibility of common genetic mechanisms (e.g Mitchell-Olds *et al.* 1995; Fokunang *et al.* 2000). Comparing individuals with diverse genetic origins, the presence of population structure can create spurious associations between unrelated phenotypes and between markers and phenotypes. The issue of population structure has received much attention in genome-wide association studies, where spurious associations between a given genetic polymorphism and a phenotypic trait can lead to false conclusions. Several approaches using linear mixed models have been developed to account for population structure and relatedness in association studies. Concurrent work in our group with a panel of 282 diverse inbred maize lines has found positive genetic correlations

for resistance to different diseases after mixed model correction for population structure and relatedness (R. Wisser *et al.* unpublished). In the analysis, population structure accounted for around 25% of the variation in the resistance phenotype. Lines originating from an environment conducive to diseases would be more likely to have resistance to several different diseases even if there was not a common genetic component, than a line originating from a non-conducive environment. In this case, the resistance phenotype could be the result of combining (stacking) many disease-specific resistance genes into a single genotype.

In segregating plant populations, MDR is also evidenced by correlated phenotypes. In this situation, if random mating is assumed, the observation of correlated phenotypes indicates either linkage or pleiotropy. In the evaluation of populations to identify quantitative trait loci (QTL) for disease resistance, the identification of co-localizing QTL for different diseases is further evidence of MDR. The phenomena of correlated resistance phenotypes and co-localizing QTL have been observed in many studies (e.g.Wisser *et al.* 2005; Wisser *et al.* 2006) though low genetic resolution has limited inference in distinguishing linkage from pleiotropy. Most studies have also examined a limited range of germplasm, limiting inference on the distribution of MDR in the species. To examine MDR in maize while addressing some of these limitations, we utilized a recently developed novel genetic resource in maize.

A maize nested association mapping (NAM) population was developed for high-resolution mapping of QTL with the possibility of gene level resolution though association mapping. NAM is a collection of 25 families each consisting of 200 recombinant inbred lines (RILs). Each of the families was derived from a cross between one of 25 diverse maize inbreds and a common

reference inbred line, B73. The resulting population of ~5,000 RILs can be jointly analyzed, providing considerable power and resolution for mapping the genes underlying quantitative trait variation. The multi-family nature of NAM permits evaluation of multiple alleles. The founding inbred parents of NAM were selected to maximize genetic diversity. As such, these lines (and the resulting progeny) display large phenotypic variation for any given trait, with quantitative disease resistance being no exception. Considering the economic importance of quantitative disease resistance and the genetic resources available in maize, we previously utilized NAM to elucidate the genetic architecture of QDR for three important maize diseases; southern leaf blight (SLB), northern leaf blight (NLB), and gray leaf spot (GLS) (Kump *et al.* in prep., Poland *et al.* in prep., Benson *et al.* in prep.). Here we further examine the genetic architecture of QDR focusing on resistance to multiple pathogens.

Fungal pathogens represent diverse and economically important taxa of plant pathogens. In maize, numerous important diseases are caused by fungi, including SLB, NLB, GLS, common and southern rust, smut, northern leaf spot, aspergillus ear rot, fusarium ear rot, and several different stalk rots. The pathogens that cause SLB, NLB, and GLS share some aspects of pathogenesis and represent three of the most economically important diseases world-wide.

SLB, caused by *Bipolaris maydis*, is a necrotrophic disease that was the cause of several devastating epidemics in the 1970s. Maize lines with T cytoplasm have extreme susceptibility to *B. maydis* Race T. In addition to cytoplasmic determinants of resistance, other nuclear genetic resistance has been identified and mapped for *B. maydis* (Balint-Kurti *et al.* 2007; Balint-Kurti *et al.* 2008). SLB is characterized by numerous small lesions that coalesce to

form large blighted areas and is prevalent in the conducive environments of the southern U.S. and the subtropics. NLB of maize is characterized by large cigar-shaped lesions caused by *Setosphaeria turcica* (anamorph *Exserohilum turcicum*). Following a short phase of biotrophic growth in the first cell infected, *S. turcica* spreads through the xylem and then grows into surrounding cells causing cell death. NLB is prevalent in cool humid environments and is detrimental in the tropical highlands. GLS is caused by the closely related fungal species *Cercospora zeae-maydis* and *Cercospora zeina*. GLS is a considerable threat to maize production in sub-Saharan Africa, and is a growing problem in the eastern U.S. with the increasing importance of agronomic practices involving minimum tillage. *C. zeae-maydis* causes small rectangular necrotrophic lesions that develop between the leaf veins.

The pathogenesis processes of *B. maydis*, *E. turcicum* and *C. zeae-maydis* exhibit both similarities and differences. Each disease is most severe and causes the most damage during the grain-fill period following anthesis. As a result, resistance is often correlated with relative maturity in these pathosystems. *E. turcicum* pathogenesis is characterized by a short period of biotrophic growth at the site of infection, followed by advance through the leaf vasculature and destructive secondary hyphal growth into surrounding cells, causing necrosis and the characteristic large cigar shaped leaf blight. In view of its long latent period, *C. zeae-maydis* is presumed to have a period of biotrophic growth, which is followed by destruction of surrounding cells. In contrast, *B.maydis* does not display any biotrophic phase, causing cell death early in pathogenesis. Unlike *E. turcicum*, *B. maydis* and *C. zeae-maydis* do not progress through the vasculature, leading to numerous small lesions

constrained by bordering veins. For each pathogen, the final stage of pathogenesis is characterized by secondary hyphal growth and corresponding host cell death.

MATERIALS AND METHODS

Plant Materials: The maize Nested Association Mapping (NAM) population consists of ~5,000 recombinant inbred lines (RILs) in 25 different families. Each RIL family is comprised of 200 F_5 inbred lines that were derived from a cross between one of 25 diverse inbred "founder" lines and the inbred reference parent B73 (Buckler *et al.* 2009; McMullen *et al.* 2009). The RILs were genotyped with 1106 SNP markers (McMullen *et al.* 2009). Multi-season trials for resistance to each of the three diseases, SLB, NLB, and GLS, were conducted as described previously (Kump *et al.* in prep, Poland *et al.* in prep, Benson *et al.* in prep). Briefly, evaluations were carried out in a single replication with lines from families grouped into a single block. An augmented design was used with incomplete blocks consisting of 20 RILs and two check lines: the respective population founder and the B73 reference parent. SLB was evaluated in North Carolina during the summers of 2006 and 2007, and in Homested, FL during the winter of 2007-2008. *B. maydis* race 0 was artificially inoculated in each trial. Disease severity was measured using a 1-9 rating scale at two time-points where 9 is no disease and 1 is completely blighted. Prior to any further analysis, the scale for SLB was inverted to correspond to the NLB and GLS ratings where low numbers correspond to low disease severity. For NLB, evaluations were conducted in Aurora, NY during the summers of 2007, 2008 and 2009. NLB trials were artificially inoculated with *E. turcicum* race 1 and evaluated at three time-points for disease severity

using a direct estimation of the percentage (0-100%) diseased leaf area. GLS was evaluated under natural infection in Blacksburg, VA during the summers of 2008 and 2009. Evaluation for GLS resistance was done at three time points using a 1-5 rating scale (Saghai Maroof *et al.* 1993).

Statistical Methods: Multivariate mixed models were used to account for field and environment effects for each disease. Model solutions provided best linear unbiased predictors (BLUPs) for each line at each rating. The line BLUPs were averaged across ratings to provide a single phenotypic value for each disease resistance. Trait distributions were then standardized to provide better comparison across diseases. These standardize values were then used for all further disease analysis. BLUPs for day to anthesis (DTA) from 4 environments over two seasons were used as trait values for relative maturity {Buckler, 2009 #1134}. Marker trait associations were evaluated using stepwise linear model selection in SAS software v9.1.2 (Proc GLMselect). Marker effects were incrementally added to the model based on the p-value of the marginal F-test. Model selection was stopped when the p-value for addition of a potential effect was $>10^{-4}$. A linear model including all selected effects was then fit and allele effects for each founder at each selected marker were calculated. Phenotypic and allele correlations were evaluated using Pearson's correlation in R statistical software (R.Development.Core.Team 2009). For phenotypic correlations, the BLUPs for each disease and for DTA were analyzed. For the disease-disease comparisons the partial correlation was also fit using DTA as a covariate. To examine correlated allele effects, the estimated allele effects from the linear model solution at a given locus for a given disease were compared to the estimated effects from a locus for a second disease where the two loci had overlapping confidence intervals.

One approach to describe the MDR value of individual lines was to calculate an MDR index. For each line, standardized phenotype values for each of the three diseases were added to make an MDR index. Only lines that had phenotypic observations for all three diseases were used. The MDR index was then treated as a single trait for identification of QTL as described above.

Confidence intervals (CI) for each QTL position were constructed by sequentially examining flanking markers. To construct the CI, the QTL marker and the first flanking marker were both fit into the full linear model. The p-value from the marginal F-test for the QTL marker was then determined. The flanking marker was considered equivalent to the QTL marker and within the 95% confidence interval if the QTL marker did not have a significant contribution to the model ($p\text{-value}<0.05$). This process was iterated by moving the flanking marker outward until the effect of the QTL marker was significant, at which point the limit of the confidence interval was declared. To compensate for regions of low marker density, imputed genotypes were examined in 1cM intervals between consecutive markers based on the expected marker class at a given genetic position.

Genome-wide nested association mapping was used to evaluate 1,606,526 SNPs for association with MDR index. SNPs for each of the 25 founder inbred lines, the reference parent B73, and Mo17 are anchored to the B73 AGP 4a.53 physical map (www.panzea.org). Missing SNP genotypes were imputed using fastPHASE software (F. Tian, P. Bradbury *et al.* unpublished) (Scheet and Stephens 2006). For an additive model, SNP markers were coded as 0 and 1 for B73 and non-B73 polymorphisms, respectively. The complete SNP dataset was then imputed to the NAM RILs.

For imputation of SNPs to RILs, pedigree and linkage information was used. SNPs that were polymorphic between B73 and a given founder were imputed to progeny based on the expected genotype at the SNP location. SNPs that were not polymorphic between B73 and a given founder were imputed as the B73 allele to all progeny of that family.

For trait-marker association, each chromosome was analyzed individually. To account for QTL on other chromosomes, phenotypic values for each chromosome were estimated by using the residual phenotypic values from a model that included all of the QTL on other chromosomes. The residual values from this reduced model were assigned to the respective chromosome and used as the phenotypic trait value from marker associations. Trait-marker association for each SNP was evaluated for each chromosome by fitting a linear model between the imputed SNP genotypes and the corresponding residual phenotypes. The strength of trait-marker association was determined from the percentile of a respective F-distribution. From the F-test, log (1/p-value) was plotted against the physical position of the SNP. Within QTL confidence intervals, the five SNP makers with the highest log (1/p) values were examined as polymorphisms in LD with potential candidate genes from the B73 filtered gene set (MaizeSequence Release 4a.53 - www.maizesequence.org).

RESULTS

Several of the inbred founders of the NAM families show MDR. With replicated entries for each founder, we obtained very accurate phenotypic assessments (BLUPs) for each of the NAM parents. The phenotypic correlation for resistance to the different diseases was positive and highly

significant, ranging from 0.62 to 0.77 (**Figure 3.1**). Several founder lines showed high levels of resistance to all three diseases (e.g. CML52 and CML247), while other lines were very susceptible to all three of the diseases (e.g. Oh7B). The NAM founders capture a large range of phenotypic variation for disease resistance and the reference parent B73 was moderately susceptible, which increases the utility of NAM for disease resistance studies. These findings were consistent with previous results based on a panel of diverse inbreds that includes the NAM founders (R. Wisser *et al.* unpublished). Wisser *et al.* also found the NAM founders were a representative cross-section of the larger more diverse panel. There was a strong negative correlation between maturity (days to anthesis, DTA) and resistance to each of the three diseases, with the strongest correlation seen between GLS resistance and DTA. This is again consistent with concurrent work on a diverse panel of inbred lines as well as previous studies (R.Wisser *et al.* unpublished). After accounting for DTA as a cofactor, the partial disease correlations for the founders were weaker. The correlation between GLS and both SLB and NLB was much lower (0.25 and 0.30, respectively), indicating that these correlations may be driven largely by relative maturity.

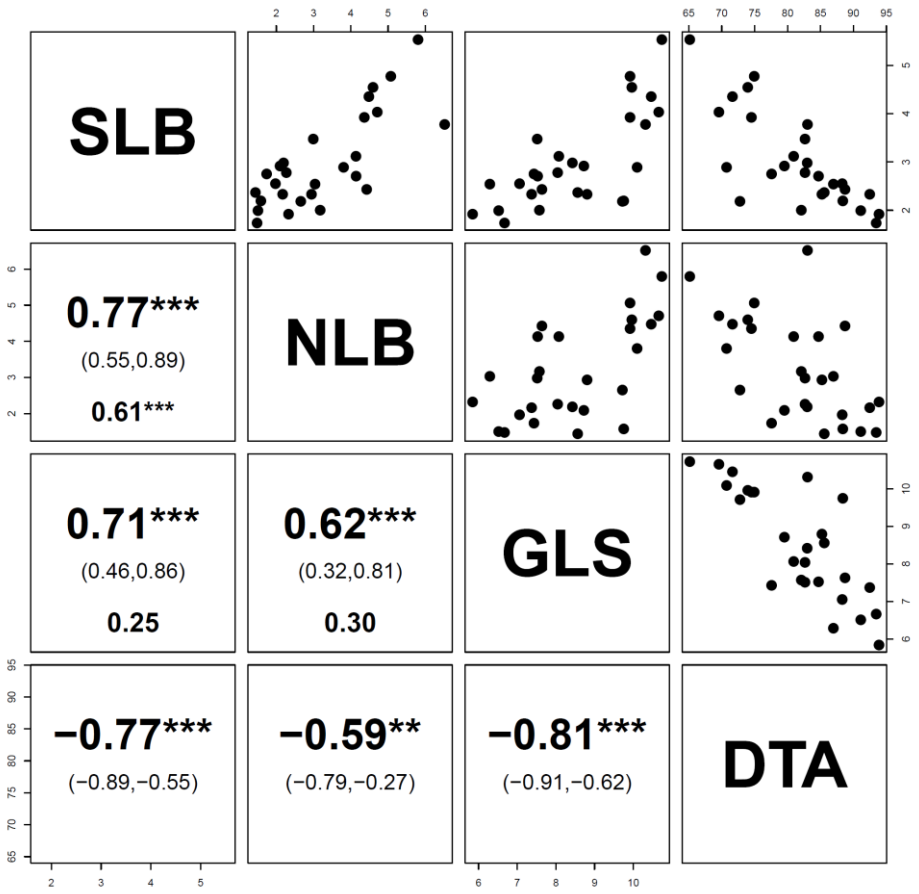


Figure 3.1. NAM founders have strong correlation for resistance to SLB, NLB, GLS and DTA.

Pearson product moment correlations for disease resistance to southern leaf blight (SLB), northern leaf blight (NLB) and gray leaf spot (GLS) in the founder inbred lines of NAM. Correlations to relative maturity are also included (days to anthesis; DTA). The upper and lower 95% confidence intervals are included in parenthesis below the correlation value. Partial correlations using DTA as a covariate are shown at the bottom. Significant correlations are shown as $p < 0.01$ ** and $p<0.001$ ***.

To begin uncovering the genetic architecture of MDR in the founders, we examined phenotypic correlation for QDR across the NAM population and within its constituent families. For NAM, there was a positive correlation between resistance for each pair of diseases, ranging from 0.42 to 0.53. The negative correlation between relative maturity and disease resistance remained, with the correlation between GLS and DTA continuing to be the strongest at -0.64 (**Figure 3.2a**). Again the partial correlation using DTA as a covariate decreased the disease-disease correlations. As the NAM lines are recombinant progeny from the founders, existing gametic phase disequilibrium has been reduced through shuffling the genome of each founder line. Unlinked QTL in a single parent have been distributed to different progeny likely contributing to a decrease in phenotypic correlations. The decrease in phenotypic correlation when comparing the correlation of founder phenotypes and that of their progeny indicated that some proportion of the MDR, as well as the association between relative maturity and disease resistance, was due to population structure or stacking of disease-specific QTL.

Figure 3.2 Correlation for resistance to SLB, NLB, and GLS in NAM.

a) Phenotypic correlations between southern leaf blight (SLB), northern leaf blight (NLB), gray leaf spot (GLS) and relative maturity (days to anthesis; DTA) in the NAM population and b) the same phenotypic correlations after removing differences between populations by subtracting each respective population mean. The upper and lower 95% confidence intervals are included in parenthesis below the correlation value. Partial correlations using DTA as a covariate are shown at the bottom.

Figure 3.2a)

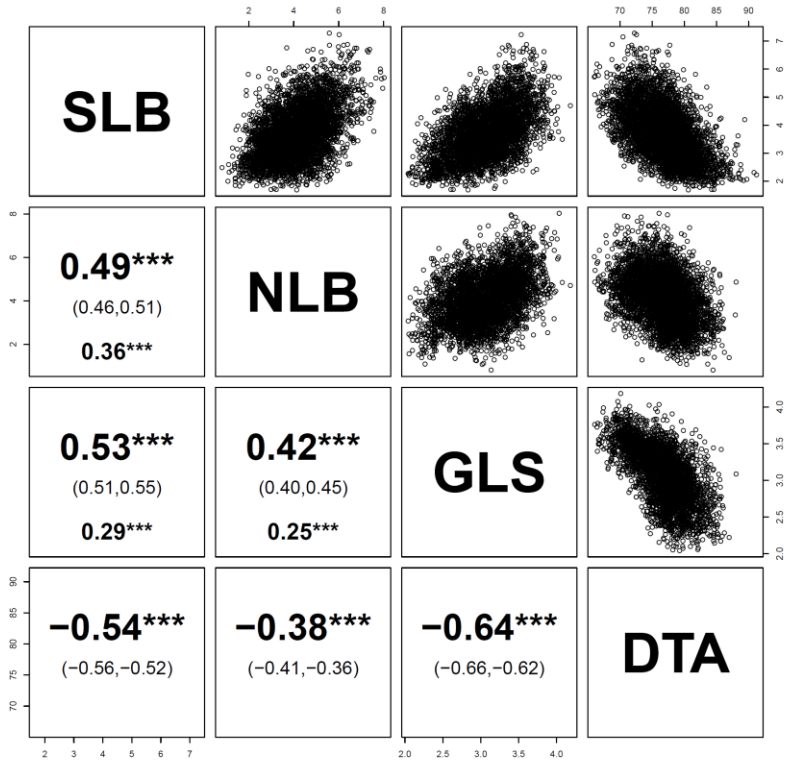
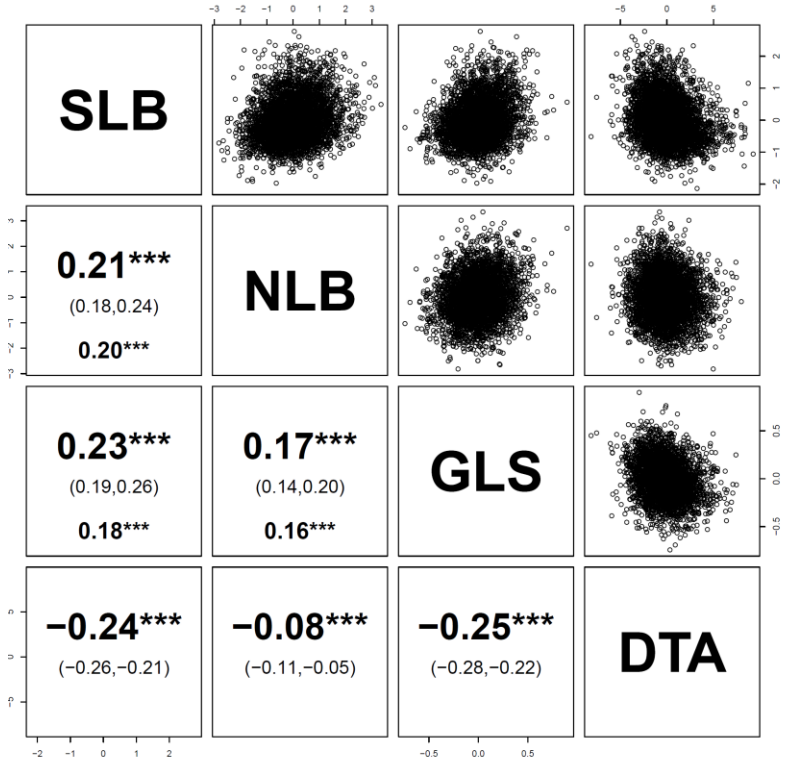


Figure 3.2b)



When comparing the whole NAM, lines within a family are more closely related than lines from different families, and different families were expected to be segregating for more or fewer QTL. To examine this in NAM, the phenotypic correlations were determined after removing the respective population means (**Figure 3.2b**). These correlation values were much lower than the analysis that included population differences. This supports the conclusion that most of the MDR observed as phenotypic correlations in the founder lines are due to pyramiding QTL rather that linkage or pleiotropic MDR genes.

To further examine the possible effect of stacking disease specific QTL in single lines, we determined the correlation between the different disease resistances within each of the NAM families. This analysis largely supported the conclusions from analyzing NAM after removing family means. When analyzing only one family, each line was derived from an independent F_2 individual and had equal probability of inheriting a given QTL. The pair-wise correlations within families were much lower than for the founders and the whole NAM population, supporting the contribution of population structure to the MDR phenotype (**Figure 3.3**).

Figure 3.3. Phenotypic correlations for individual NAM families

Phenotypic correlations between southern leaf blight (SLB), northern leaf blight (NLB), gray leaf spot (GLS) and relative maturity (days to anthesis; DTA) for individual families of NAM. Significant correlations are referenced by *; **; and *** for p-value < 0.05, <0.01, and < 0.001, respectively. The upper and lower 95% confidence intervals are included in parenthesis below the correlation value. Partial correlations with DTA as a covariate are shown at the bottom for disease-disease comparisons.

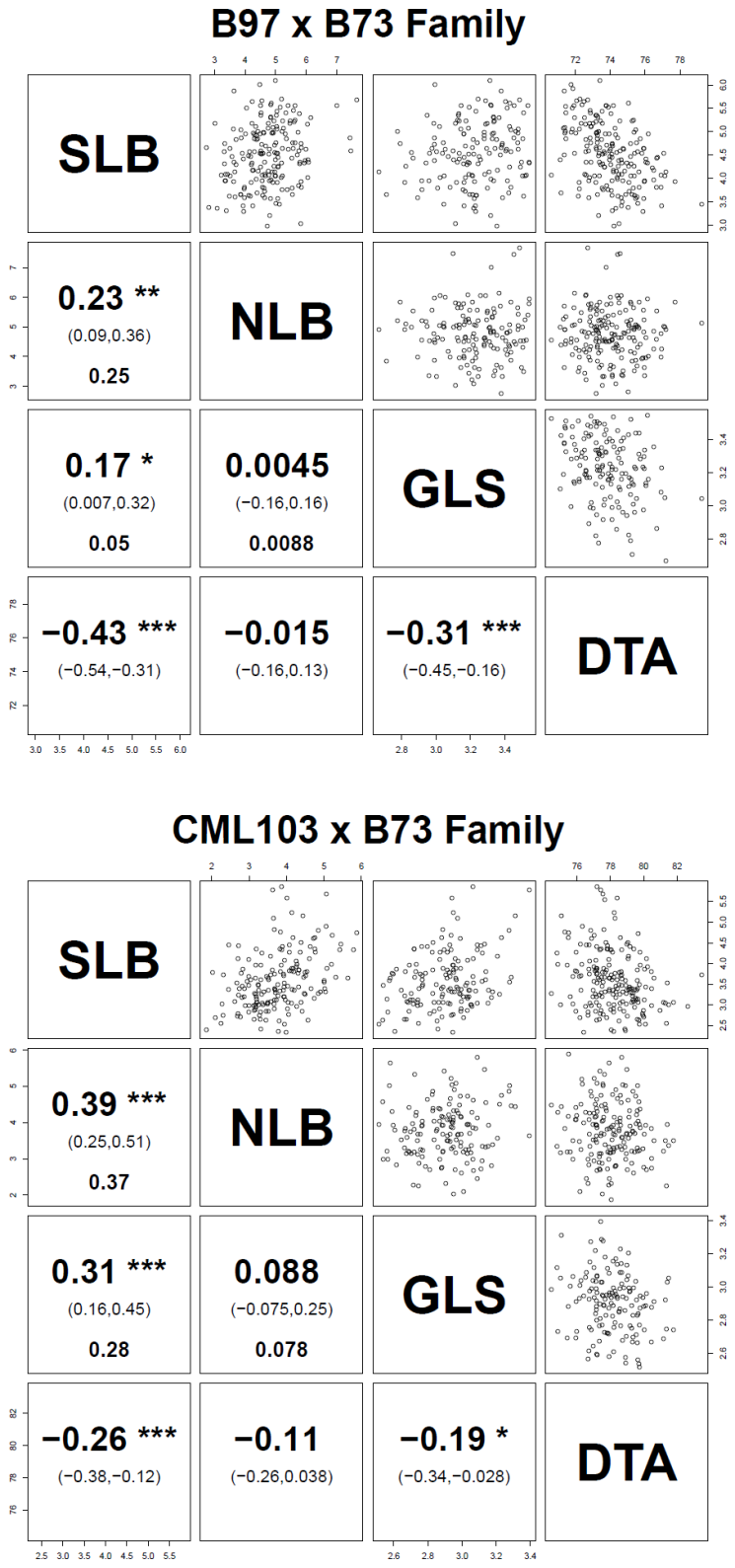
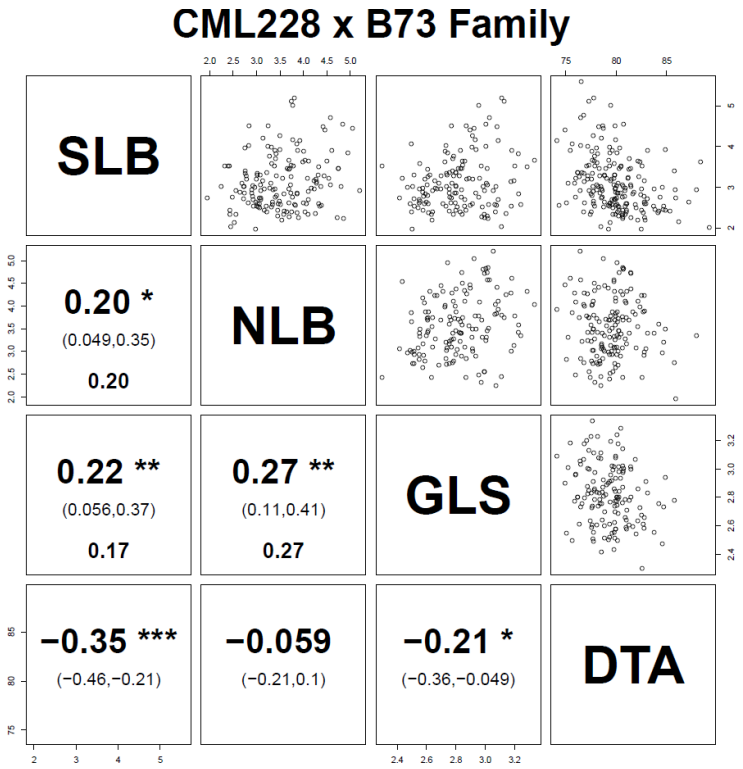


Figure 3.3 Phenotypic correlations for individual NAM families

Figure 3.3 (Continued)



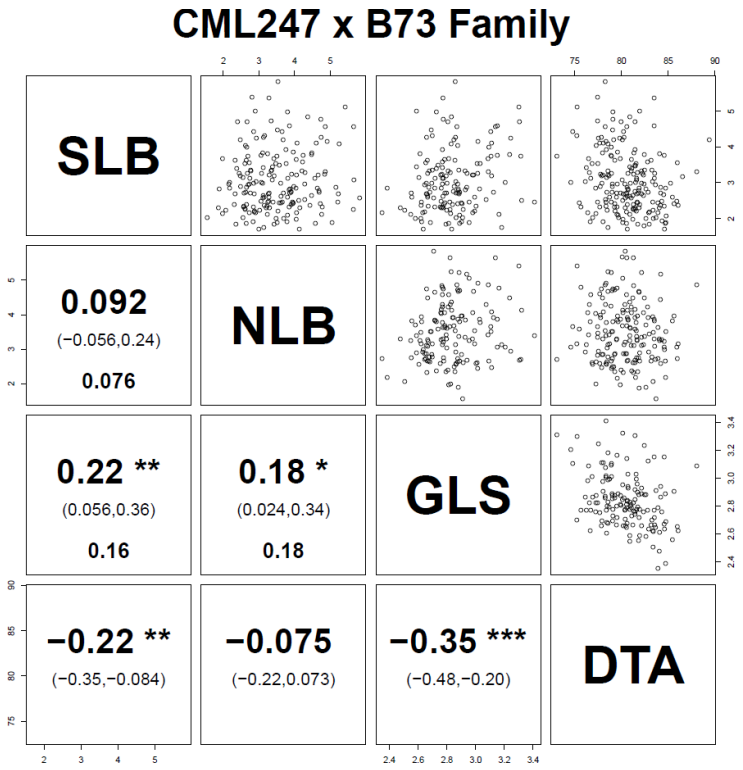
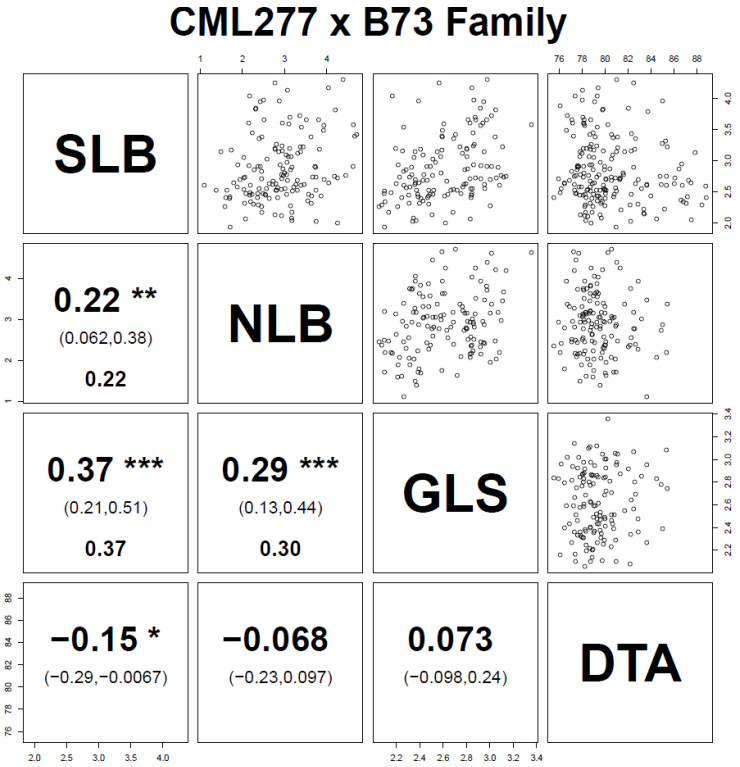


Figure 3.3 (Continued)



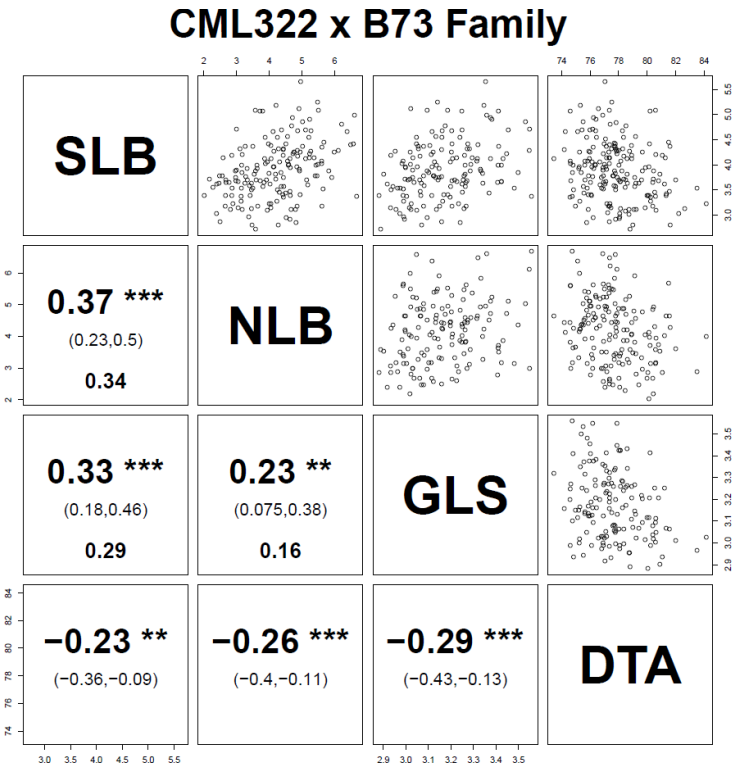
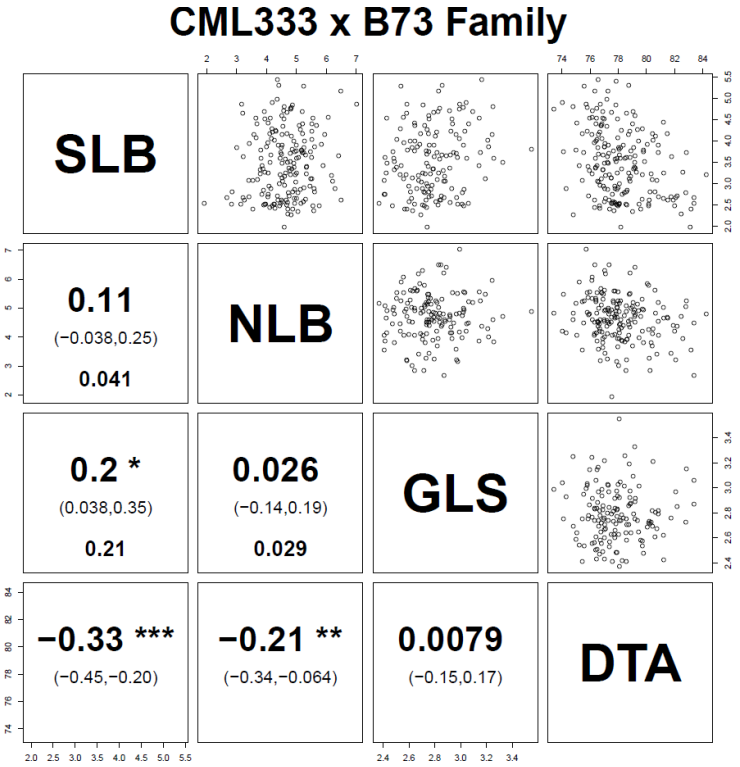


Figure 3.3 (Continued)



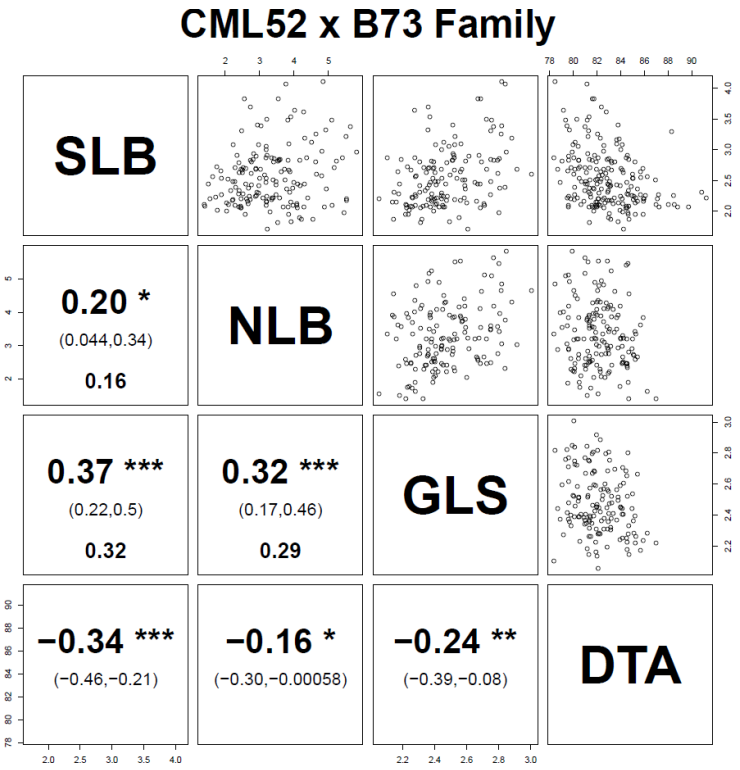
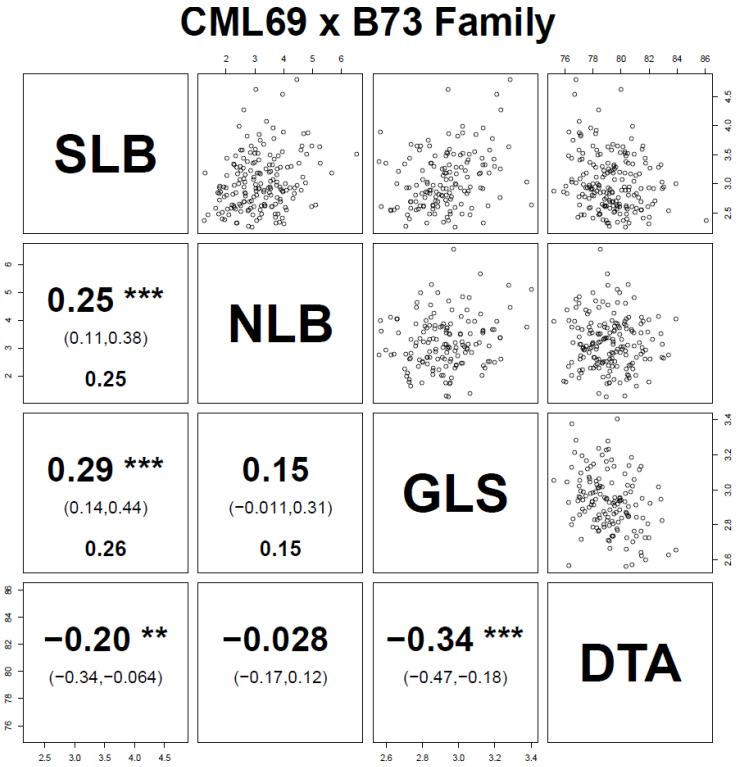


Figure 3.3 (Continued)



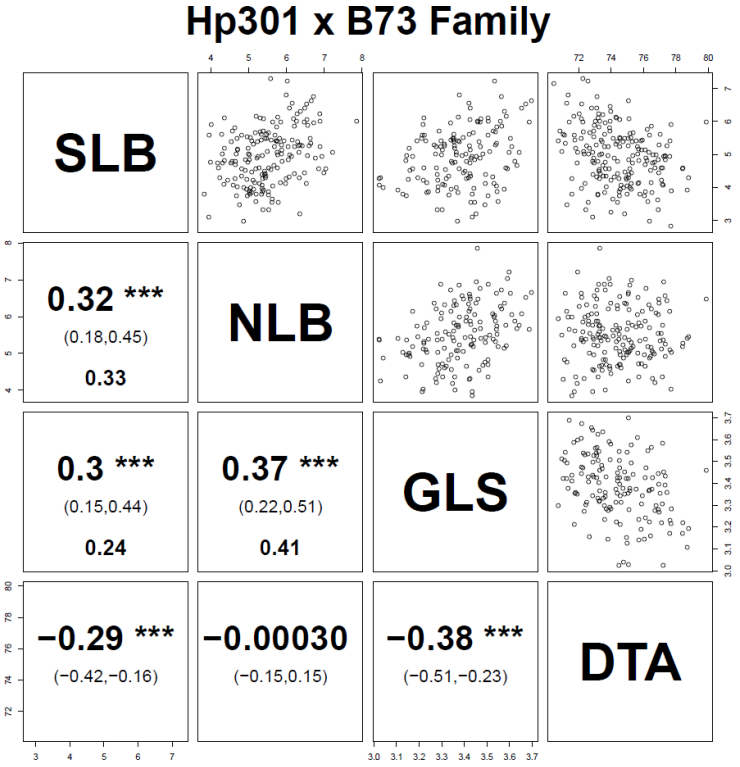
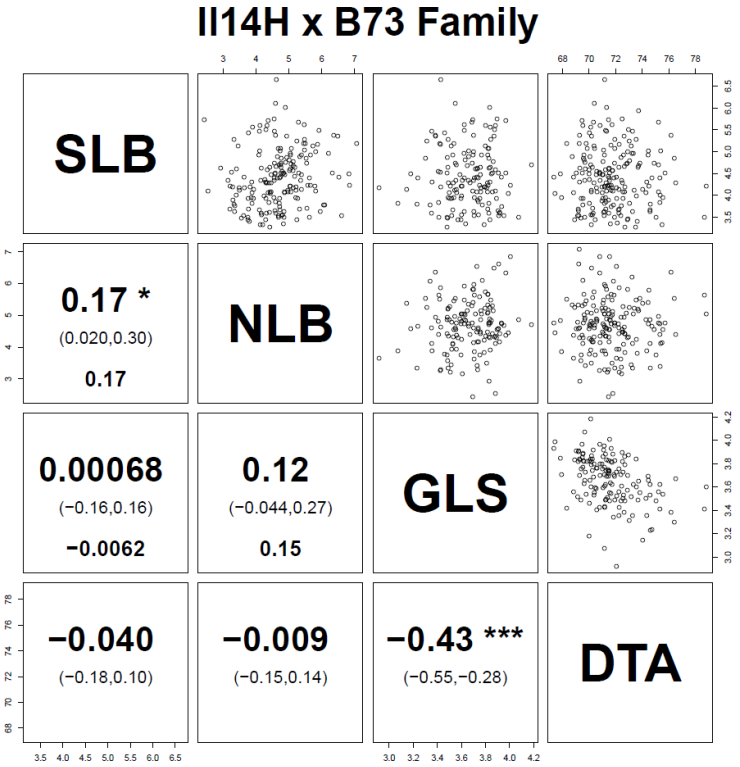


Figure 3.3 (Continued)



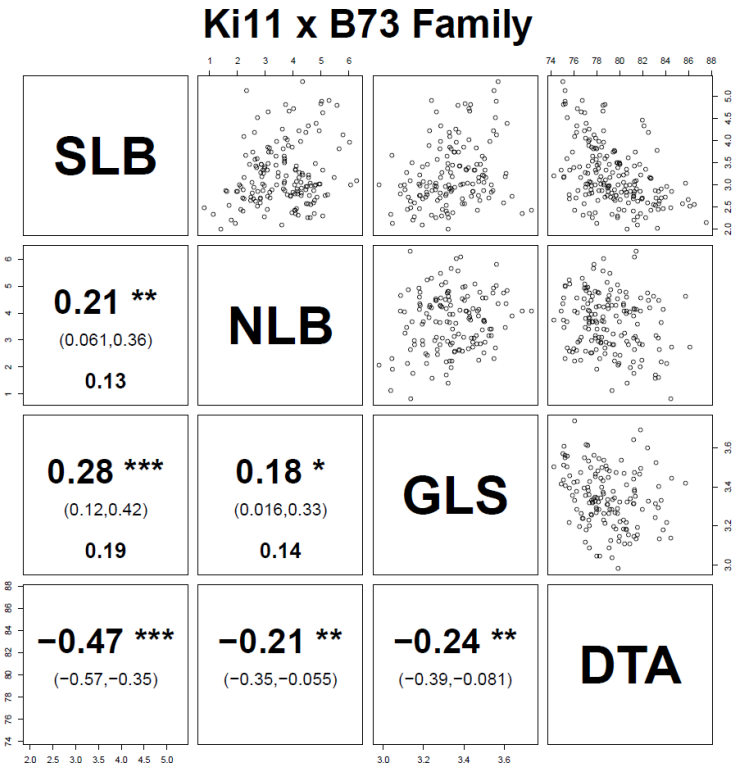
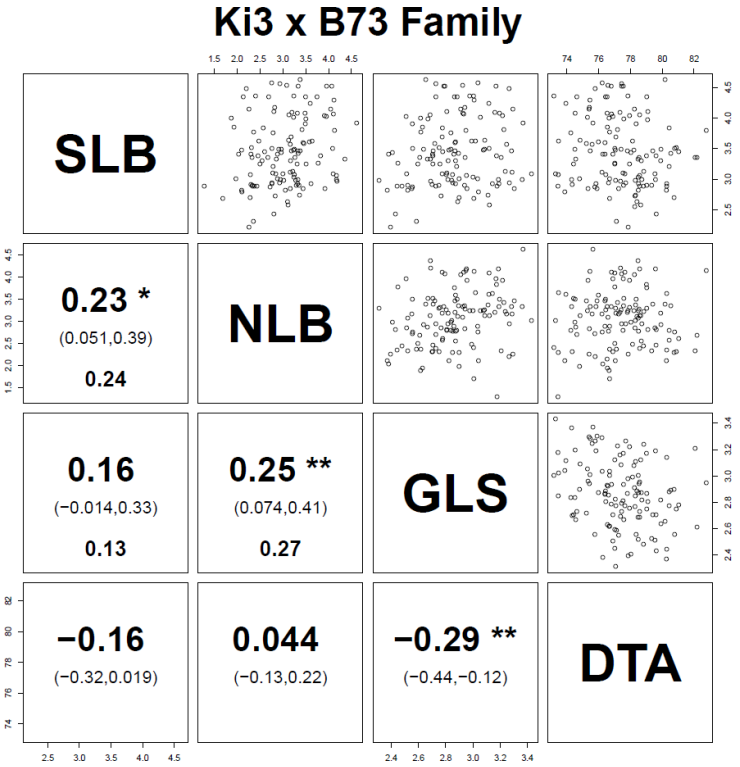


Figure 3.3 (Continued)



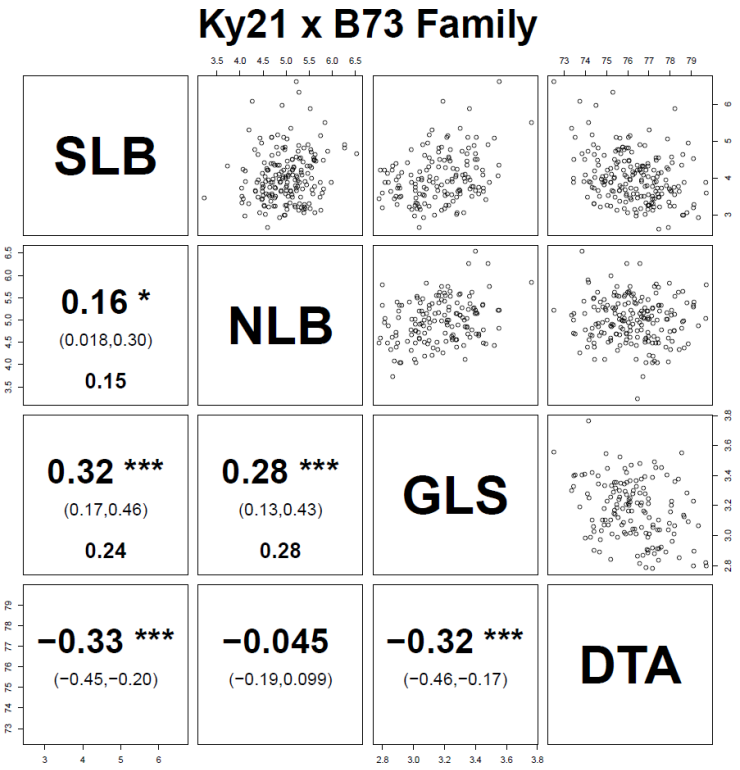


Figure 3.3 (Continued)

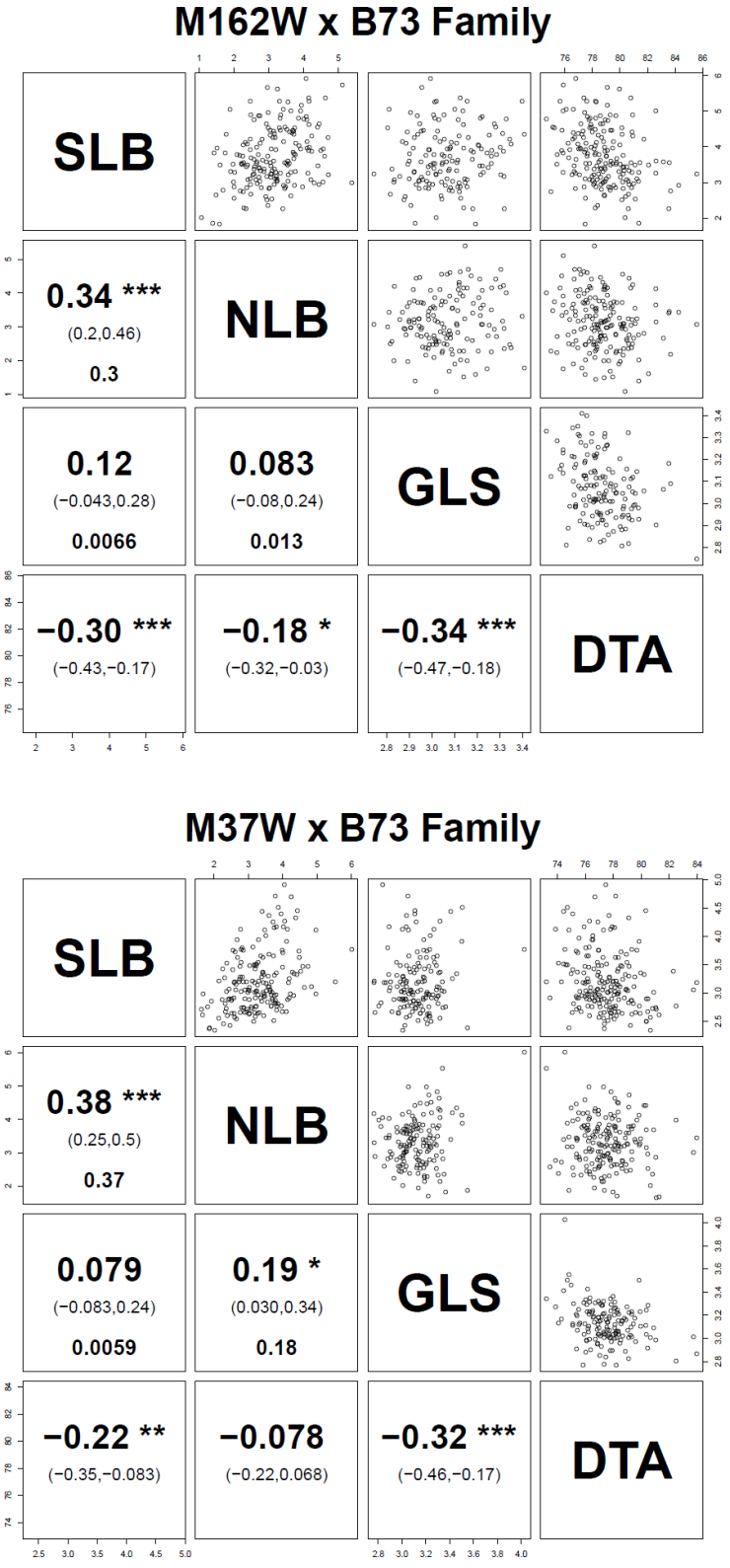
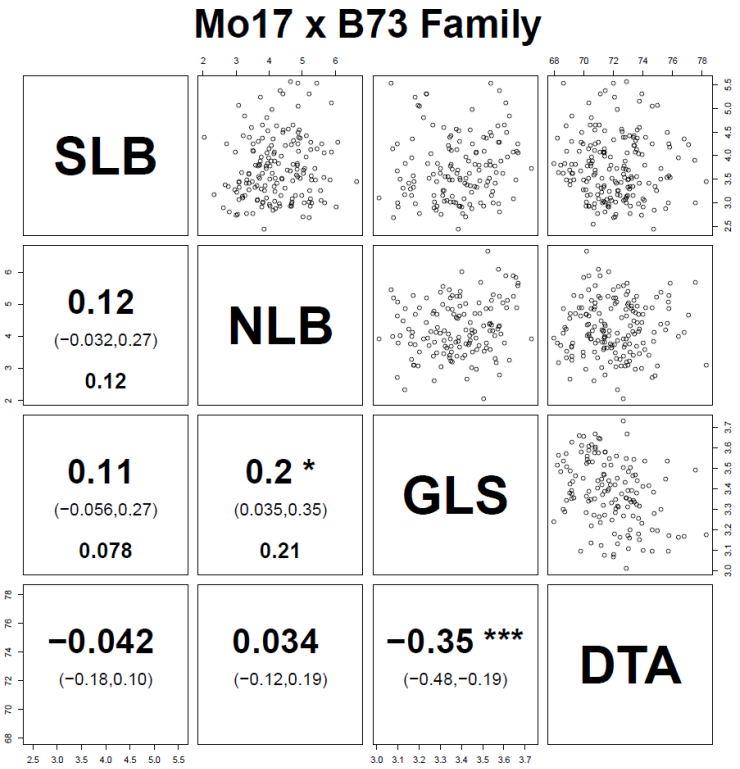


Figure 3.3 (Continued)



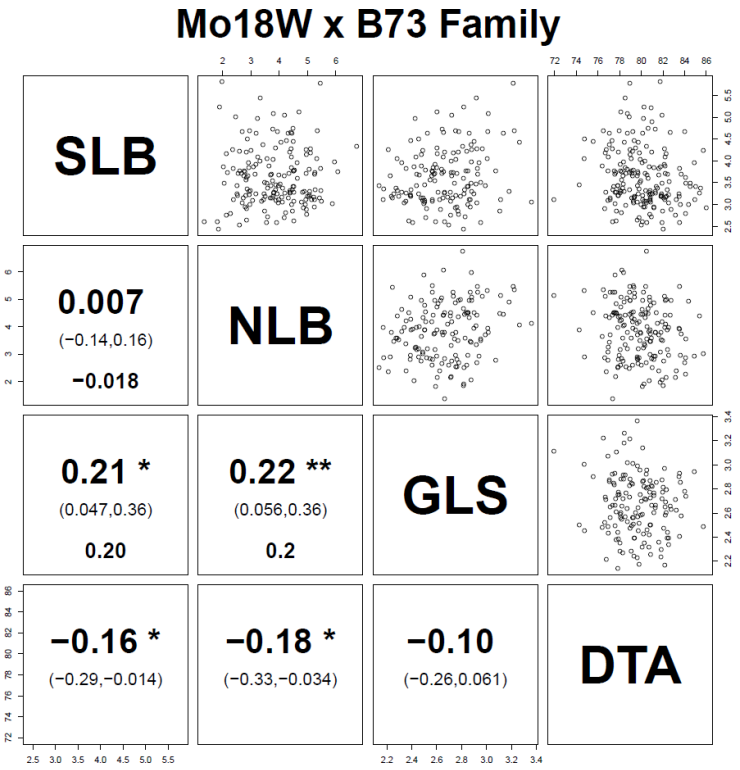
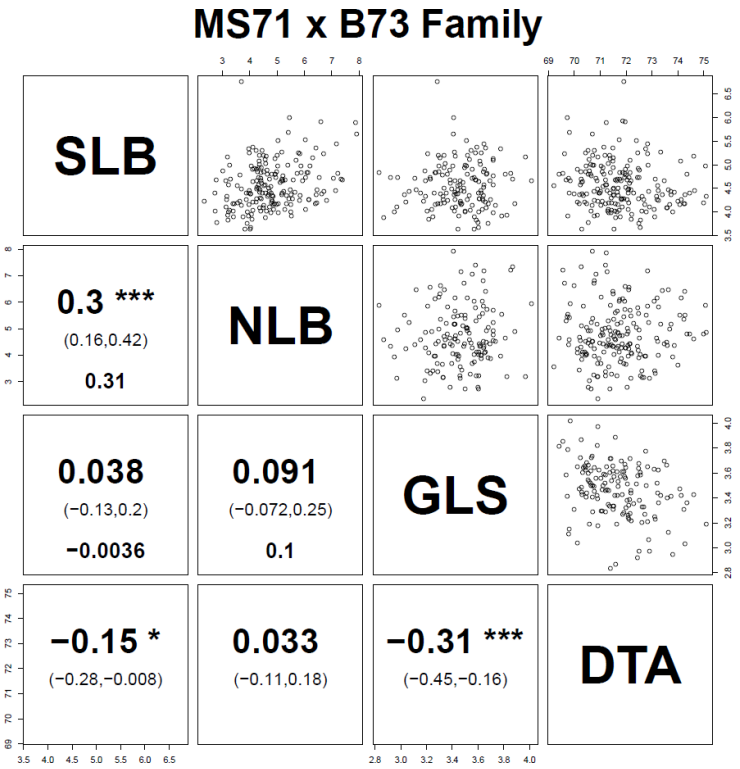


Figure 3.3 (Continued)



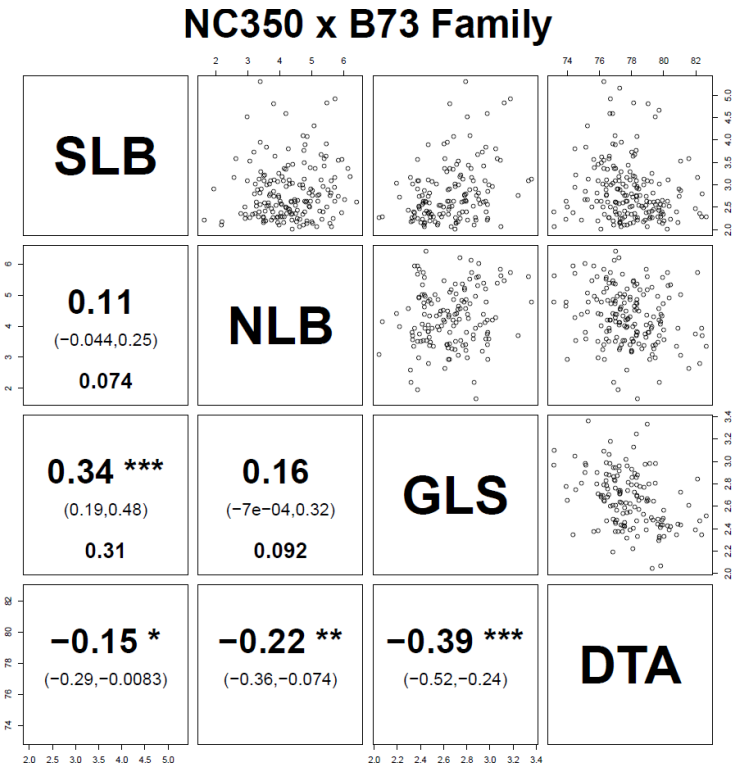
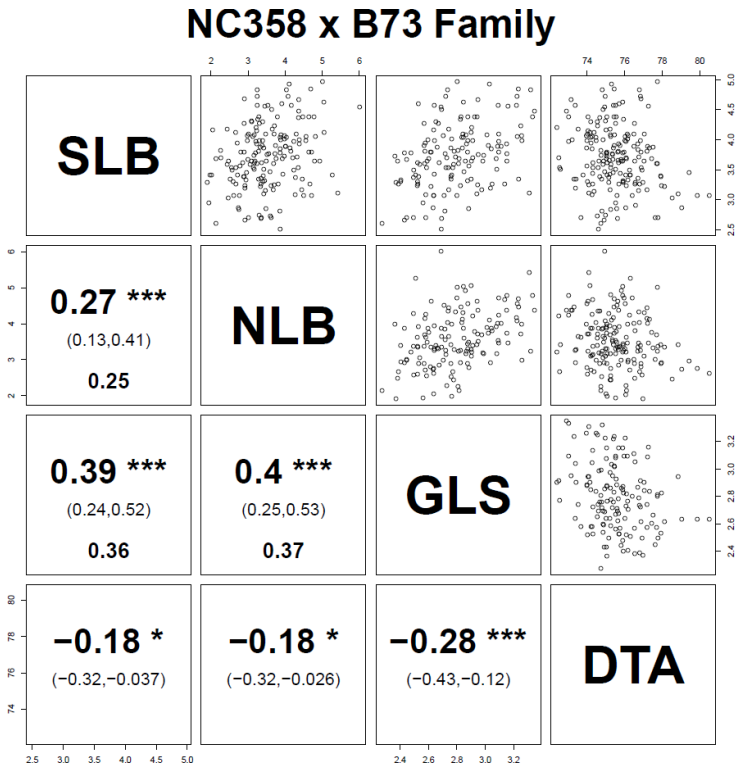


Figure 3.3 (Continued)



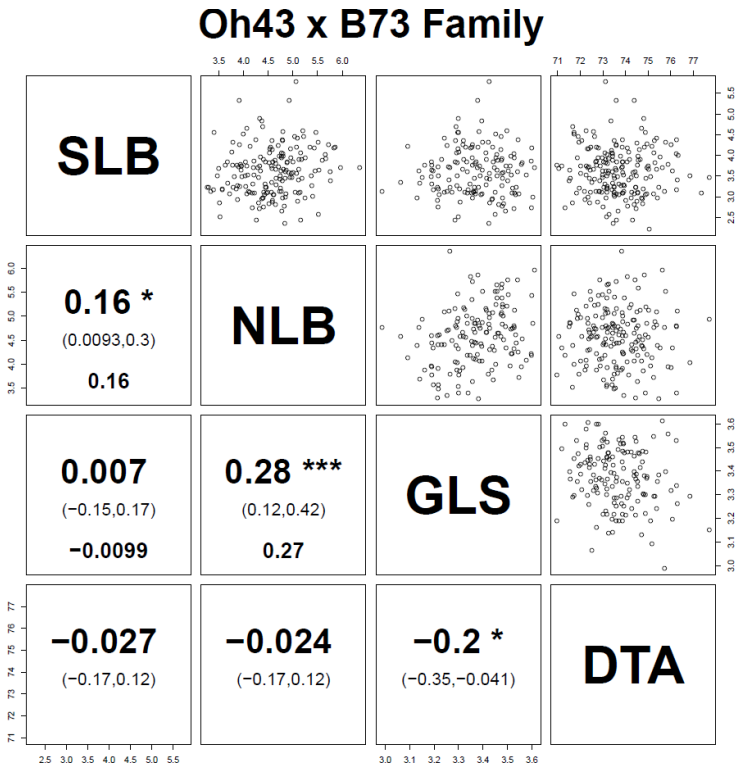
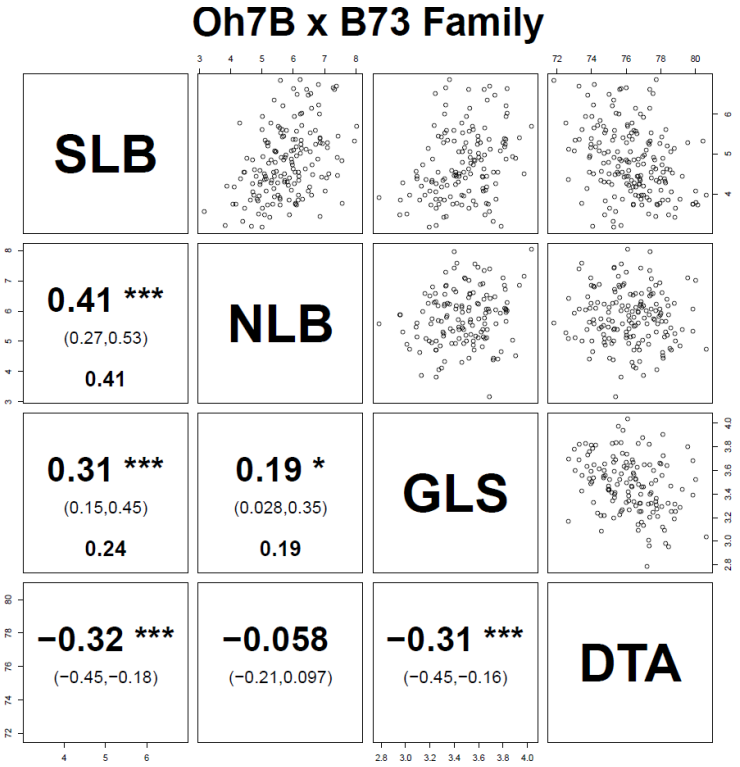


Figure 3.3 (Continued)



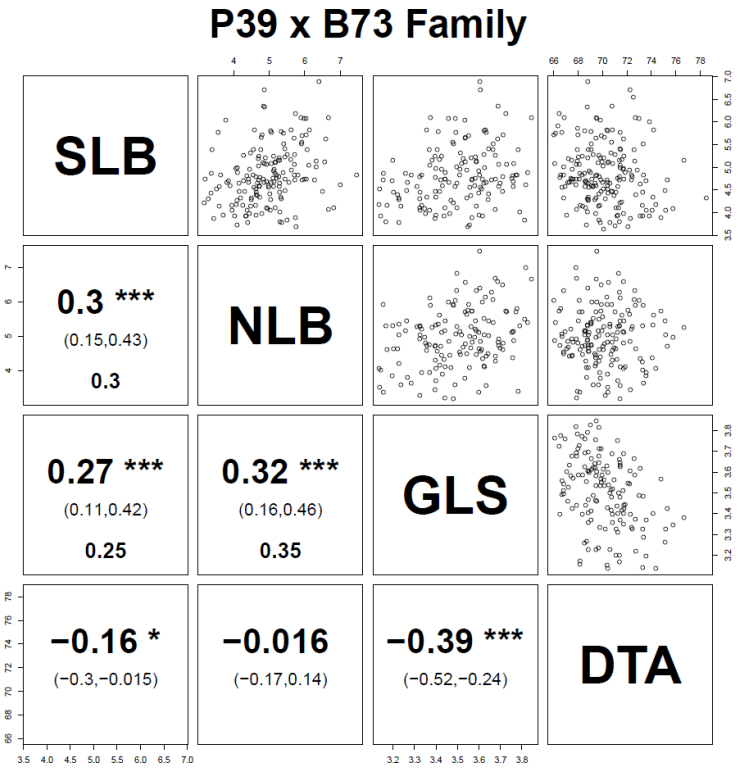
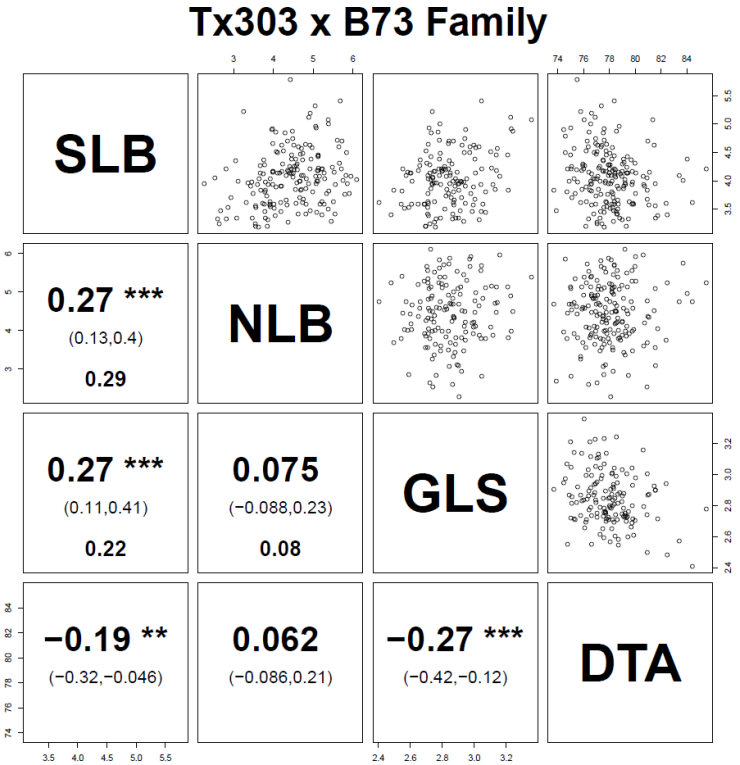
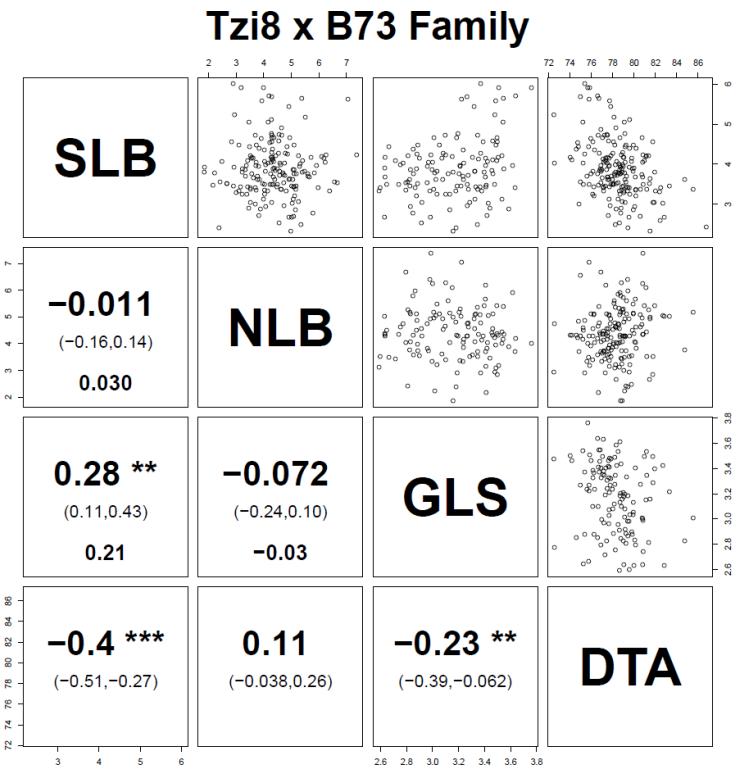


Figure 3.3 (Continued)





The significant within-family correlations ranged from 0.16 to 0.41. Not all of the disease correlations in a family were significant; of the 78 within-family pair-wise correlations (from 26 families with 3 pair-wise correlations each), 55 were significant (71%). Based on the confidence intervals of the NAM correlations, only one of the within-family MDR correlations was not significantly lower (95% CI) than the respective NAM correlation. This indicates that uneven distribution of QTL among the NAM founders (and derived progeny) contributes to the MDR phenotype. The MDR phenotype, however, was detected even with the most conservative reduction of population structure through examination of independently derived lines within a single family (e.g. no population structure).

Though lower than that seen in the founders, much of the association of relative maturity and disease resistance remained within single families. The significant correlations ranged from -0.16 to -0.47, with 43 of the 78 relative maturity-disease resistance correlations being significant. Of the 43 significant within-family correlations, 23 were between GLS and DTA, confirming the strong effect of relative maturity on GLS resistance. The reduction in relative maturity and disease resistance association also indicates that some of the correlation in the founders can be attributed to population structure.

The number of QTL and the magnitude of the QTL effects in a line presumably determines, to a large extent, the overall phenotype of that line. To examine this effect in the NAM founders, the sums of the estimated allele effects for each disease were compared. As with the phenotypes of the founders, there was a strong positive correlation between the predicted phenotype for each of the diseases (**Figure 3.4**). These correlations were based on an additive genetic model for disease resistance and were slightly

(though not significantly) higher than the phenotypic correlations. These estimates were derived from examination of almost 200 progeny from each parent and therefore give a very precise estimate of the additive genetic value for that line. Most of the phenotypic variance (and covariance) is accounted for in the additive genetic prediction model as the correlations were as high as the phenotypic correlations.

In segregating bi-parental populations, quantitative traits can be separated into single loci with individual effects. When considering pleiotropy, this allows the comparison of individual loci for several traits of interest. Co-localization of QTL has been observed in previous mapping studies and meta-analysis of QDR (Wisser *et al.* 2005; Wisser *et al.* 2006). Though NAM has higher resolution than earlier mapping studies, the resolution remains insufficient to distinguish between linkage and pleiotropy, which would require gene-level resolution.

Genomic regions of potential MDR were identified at the positions where the confidence interval of two or more QTL overlapped. Using this criterion, we identified 23 chromosomal segments that contained QTL for two or more diseases (**Figure 3.5**). Seven of these locations had QTL co-localizing for all three diseases. We compared the position of QTL for days to anthesis (DTA) (Buckler *et al.* 2009) to the QTL for each disease. We found 28 QTL for DTA co-localized with one or more disease QTL.

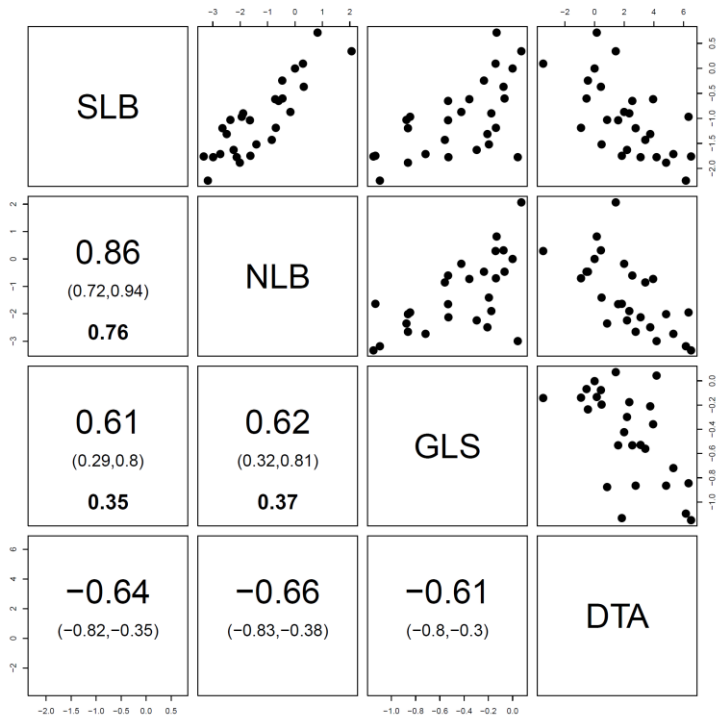
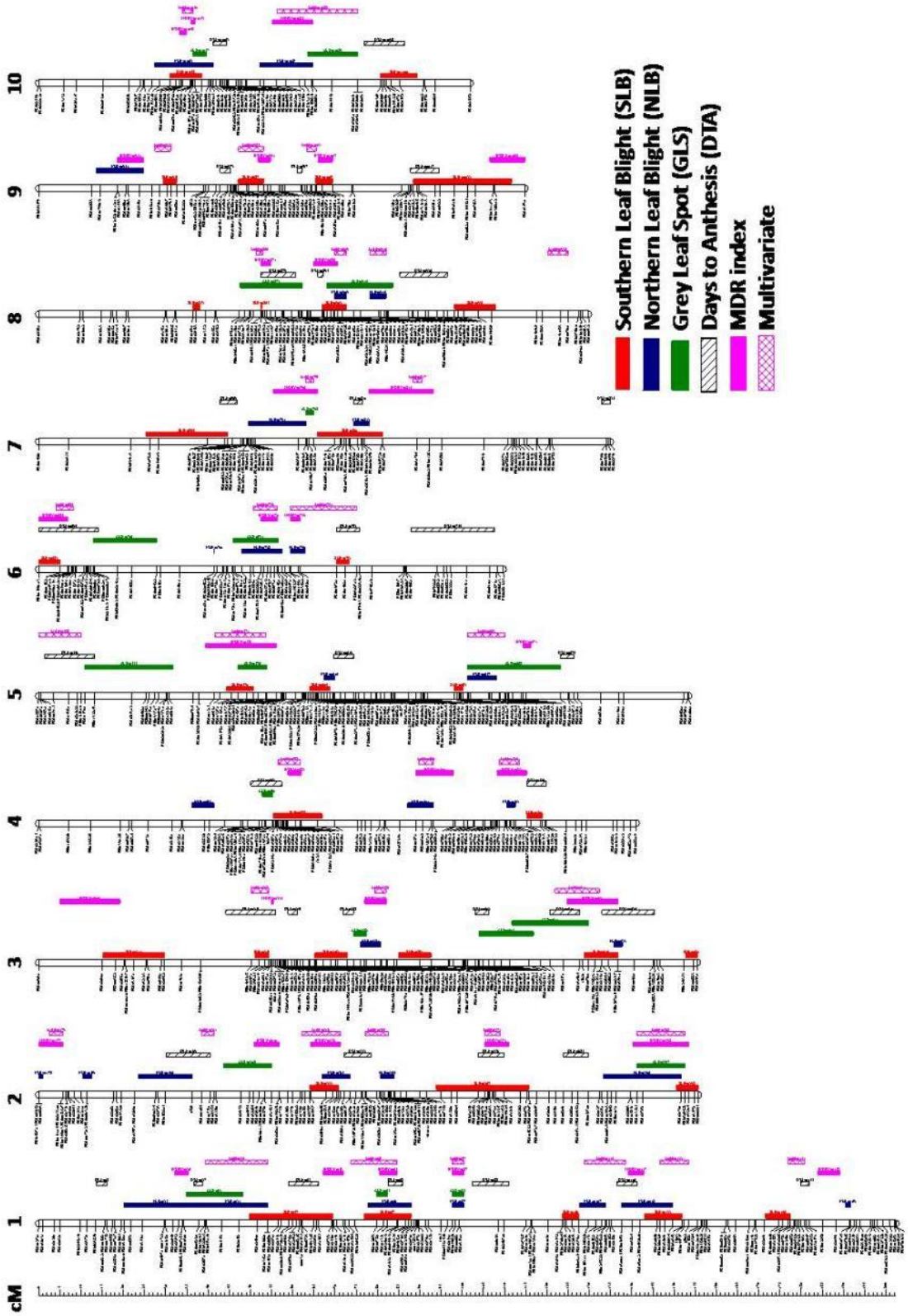


Figure 3.4 Predicted phenotypes for NAM founders are correlated. Correlation of sum of all allele effects for each NAM founder for southern leaf blight (SLB), northern leaf blight (NLB), gray leaf spot (GLS) and relative maturity (days to anthesis; DTA).

Figure 3.5. Genetic linkage map of NAM which identified QTL for SLB, NLB, and GLS and QTL for DTA

Consensus linkage map of NAM families showing the genomic position of quantitative trait loci (QTL) for southern leaf blight (SLB), northern leaf blight (NLB), gray leaf spot (GLS) resistance and QTL for relative maturity (days to anthesis; DTA). 95% confidence intervals for QTL are shown. The allele effects for QTL that had overlapping confidence intervals were compared for evidence of multiple disease resistance.


Southern Leaf Blight (SLB)
Northern Leaf Blight (NLB)
Grey Leaf Spot (GLS)
Days to Anthesis (DTA)
MDR index
Multivariate
10
9
8
7
6
5
4
3
2
1
cM

Reasoning that pleiotropic genes underlying co-localizing QTL would be more likely to have correlated allele effects than linked disease specific genes, we then examined the correlation of allele effects at co-localizing QTL to determine the possibility of pleiotropic genes underlying the MDR phenotype in NAM. This is possible as the multi-parental design of NAM allows comparison of different alleles from each of the founder inbreds. At seven locations, there was a significant positive correlation in QTL allele effects. Three of these QTL positions had correlated effects for NLB and GLS resistance and three for NLB and SLB. At one location on Chr.3 (77cM), allelic effects for all three diseases were correlated (**Figure 3.6**). A QTL for DTA was located in the region and also had correlated allele effects. The allele effects for disease-disease correlations were positively correlated, while the allele effects for DTA-disease were negatively correlated. This corresponds to the observed phenotypic correlations. Although the allele effects were correlated at this locus, the confidence intervals for SLB and NLB QTL did not overlap, indicating that the 95% confidence intervals may be too conservative.

Three additional QTL had correlated allele effects for SLB and NLB, one on Chr.2 at 78-cM, a second on Chr. 5 at 70-cM, and the third on Chr.7 at 76.5 cM. Three additional QTL also had correlated allele effects for NLB and GLS. These QTL were located on Chr. 1 at 61.6 and 84.9-cM and on Chr. 10 at 64.8 cM. At each location, the allele effects were positively correlated in agreement with the observed phenotypic correlations. The strongest allele correlations were found on Chr. 3 at the location where QTL for all three diseases are located.

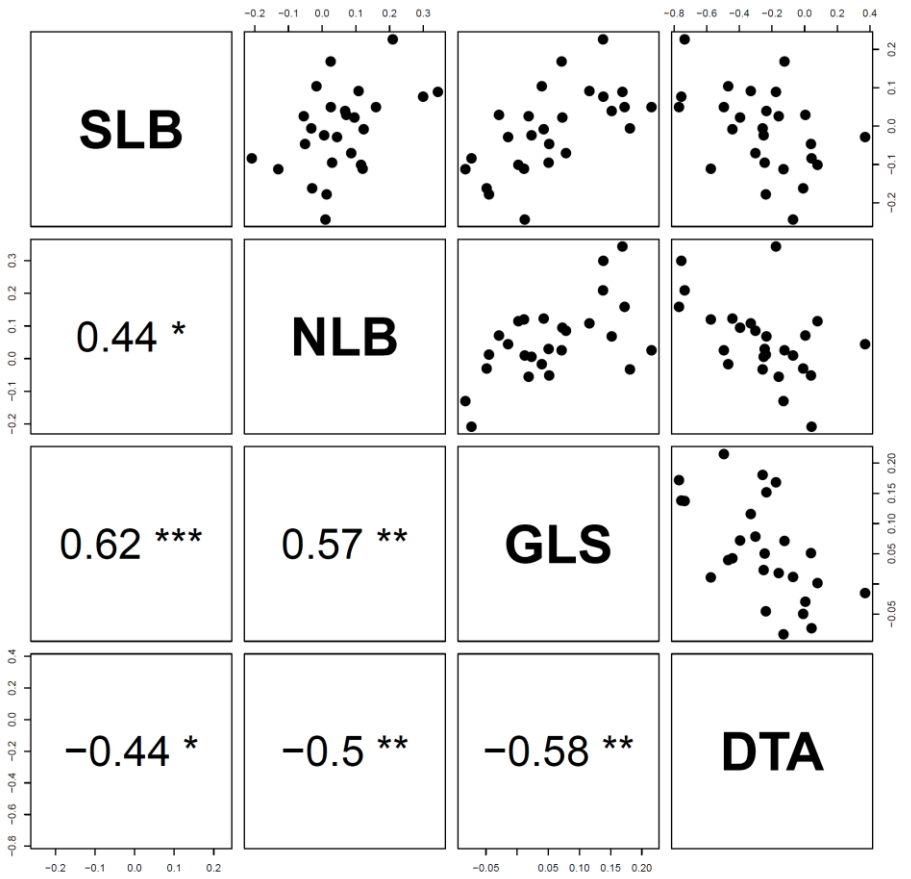


Figure 3.6 Estimated allele effects for the NAM founders for a QTL/QTL on Chr. 3 (65-80cM) are correlated

Correlation between estimated allele effects for each NAM founder for southern leaf blight (SLB), northern leaf blight (NLB), gray leaf spot (GLS) and relative maturity (days to anthesis; DTA). Correlations are significant at p-value< 0.05*, 0.01** and 0.001***.

Not including the locus on Chr. 3 where all four traits had correlated allele effects, there were six additional loci with negative correlations between SLB and DTA and one additional locus with negatively correlated effects for GLS and DTA. The correlations ranged from -0.39 to -0.44 for SLB-DTA and the correlation was -0.44 for GLS. The strongest disease-maturity correlation of -0.58 was between GLS and DTA at the locus on Chr. 3 where all allele effects were correlated. We did not observe any significant allele correlations that contrasted to the phenotypic correlations (i.e. negative for disease-disease or positive for disease-maturity).

To explore the possibility of genetic variability for all three diseases, an MDR index was calculated by adding standardized phenotypic values for each of the three diseases. For the MDR index, 34 QTL were identified at $p < 10^{-4}$ (**Figure 3.5)** . Several of these QTL clearly correspond to locations were several QTL for individual disease are also located (e.g. Chr. 1 at 85-cM and 100-cM.) At these locations, it appears that the MDR index is capturing the underlying genetics of MDR. However, there are several locations where QTL for two or more diseases co-localize but where no MDR index QTL were identified (e.g. Chr. 10 at 35-40cM). Other QTL identified for the MDR index appear to position between two adjacent single disease QTL (e.g. Chr. 10 at 62-cM). Finally, it is apparent at several locations that the MDR index was driven by variability for only one trait. This is clearly seen on Chr. 9, where three MDR index QTL are at the same position as three QTL for SLB, and there are no QTL for the other two diseases in the region.

We used a multivariate model as a second approach to directly capture genetic variability for MDR. Similar to the univariate model, stepwise model selection was performed using the p-value from Wilk's F-test to determine

inclusion of a marker effect using all three disease traits as the multivariate response. Multivariate stepwise selection identified 31 QTL at a selection threshold of p-value < 10^{-6}. There was general agreement between the multivariate model and univariate model for MDR index, though many of the QTL identified with the multivariate approach did not correspond with the MDR index. Of the 31 multivariate QTL identifed, 23 were overlapping with QTL for the MDR index.

Genome-wide nested association was used to evaluate >1.6M SNPs for association with the MDR index. At each QTL location, the five SNPs with the most significant association tests [log (1/p) from F-test] were examined as polymorphisms in close proximity (<10-kb) to potential candidate genes for MDR. Almost 200 QTL-associated genes were identified at the SNP positions examined. Several intriguing associations were seen: a small proportion of the genes identified were recognized as having a known role in pathogen defense. Many potential MDR genes identified did not have any known function.

On Chr. 1 (163.86 Mb) a germin-like gene (among others) was associated with the MDR index. In rice, a family of germin-like genes was associated with quantitative disease resistance to two distinct pathogens (Manosalva *et al.* 2009). A gene with kinase and leucine-rich repeat domains was associated with a MDR index QTL on Chr. 9. This type of gene is known as a receptor-like kinase (RLK) and has been associated with both gene-for-gene resistance (Wang *et al.* 1998) and basal resistance (Gomez-Gomez and Boller 2000; Wan *et al.* 2008). Four other genes with no known connection to disease resistance were also associated with SNPs at this QTL interval. A pathogenesis-related ethylene response factor (ERF) was identified on Chr. 3

at 165.23 Mb. ERFs are known to be involved with initiation of defense mechanisms in response to ethylene and jasmonic acid signaling.

DISCUSSION

The 27 diverse inbred founders of NAM have correlated levels of quantitative resistance to three important maize diseases, southern leaf blight, northern leaf blight and gray leaf spot. The inbred lines on the resistant side of this spectrum can be considered to have multiple disease resistance (MDR). The underlying genetic architecture of MDR in the NAM parents is likely due to 1) combination of multiple unlinked disease specific genes, 2) clusters of linked disease specific genes, 3) pleiotropic genes conditioning MDR, or 4) any combination of the above. We utilized large segregating populations derived from these founder inbreds to examine the underlying genetics of the MDR phenotype. We incrementally dissected the MDR phenotype by examining phenotypic correlations with decreasing levels of population structure. Starting with the founder inbred lines, we found a strong correlation among resistances to SLB, NLB and GLS as well as relative maturity (DTA). We then advanced to analysis of recombinant progeny across families, to progeny within families, and finally to individual chromosome segments. With each progressive step there was a concurrent reduction in the MDR phenotype. This indicates that much of the MDR phenotype found in the NAM parents is a result of stacking disease-specific QTL.

We found evidence of co-localizing QTL for the three diseases in this study. Although the very large population size of NAM allows high resolution mapping of QTL, gene based resolution through linkage mapping is still out of reach, limiting our ability to distinguish between linkage and pleiotropy.

Regardless, we found several instances of co-localizing QTL for resistance to different diseases. In maize, days to anthesis (male flowering) and days to silk (female flowering) are highly correlated traits. In NAM, most QTL for these two traits co-localize and all have correlated allele effects supporting the hypothesis that a set of pleiotropic genes controls most of the variance in both traits (Buckler *et al.* 2009). One caveat of this approach is that there are documented cases where a single gene can increase resistance to one pathogen while increasing susceptibility to another (e.g.Nagy *et al.* 2007). This would lead to negatively or non-correlated allele effects even though a single gene is affecting resistance to both pathogens. Where this resistance/susceptibility tradeoff has been observed the negative pleiotropy has corresponded to different pathogen lifestyles (e.g. necrotrophic/biotrophic). Working with three necrotrophic pathogens as we are in this study would hopefully limit similar confounding results.

To examine the evidence of MDR at co-localizing QTL, we compared the estimated allele effects of the 27 founder lines. Among the 23 QTL examined, there were seven loci with correlated allele effects. Three of these QTL had correlated allele effects for NLB and GLS, while three had correlated effects for NLB and SLB. One of these QTL had positively correlated allele effects for all three diseases. At the same locus, flowering time allele effects were negatively correlated with resistance alleles. The correlation of estimated allele effects of the founder inbred lines lends support to the hypothesis of pleiotropic genes for MDR. These “MDR genes” could be genes that are involved in general plant defense resulting in resistance to a diverse array of pathogens. They could alternatively be genes with primary roles in plant development that secondarily affect pathogenesis (Poland *et al.* 2009).

As expected from previous studies, there was a negative correlation of relative maturity with disease severity in the pathosystems examined, with later-maturing lines being more resistant. As the founder lines originated from very different breeding programs ranging from tropical to temperate, the correlation could be due to population structure as tropical lines (later maturing) would also have been under higher selection pressure for disease resistance. As with the disease correlations, the association of relative maturity and disease resistance seen in the founder inbreds decreased with evaluation of recombinant progeny. However, significant correlations remained. Examination of co-localizing disease QTL and maturity QTL identified 23 locations where disease and maturity QTL overlapped. Six of these locations showed significant negative correlations (consistent with the negative phenotypic correlation) between maturity and disease QTL.

In an attempt to capture genetic variability for multiple disease resistance, an index was calculated for the phenotypic values of the three diseases in this study. The MDR index is easy to interpret phenotypically as it is simply the sum of the individual disease values. Lines with a low MDR index score would be considered to have resistance to all three diseases. However, the interpretation of QTL identified for the MDR index is less clear. In some cases (particularly Chr. 9) it is apparent that the loci identified are for resistance to only a single disease. This is likely due to a lack of QTL for other diseases in the region. In other cases, the QTL for MDR index are at the same positions as several disease QTL. At these positions the MDR index clearly appears to be capturing any genetic component of MDR. These contrasting results pose difficulties in using the MDR index to identify regions with MDR effect as the QTL identified might only be reflecting genetic

resistance for a single disease.

For a second approach to modeling genetic variance for MDR, multivariate model selection was used to identify loci conditioning an effect on all three disease traits. As with the MDR index, interpretation of QTL identified by multivariate models is difficult. The MDR index and the multivariate model were comparable strategies for mapping MDR QTL. However, there were discrepancies between the two models. In general the MDR index appeared to have greater consensus with the individual trait QTL mapping results, though it is difficult to determine if this is a true representation of MDR.

To identify potential MDR candidate genes, genome-wide nested association mapping was evaluated using >1.6M SNPs from the maize hap-map. SNPs were imputed to the NAM RILs and association tests were conducted for each of the SNPs using the MDR index. Three predicted genes that have previously been associated with plant disease resistance were associated with SNPs in the MDR index QTL intervals. Little inference could be made on the remaining associated genes. The defense related genes can be considered leading candidate genes, but it would be unwise to exclude the other genes, especially in light of recent findings of apparently unrelated genes being involved in QDR (Fukuoka *et al.* 2009; Krattinger *et al.* 2009). It is also important to recognize that functional polymorphisms for QTL could be distant from the actual defense related genes (Salvi *et al.* 2007).

Based on this analysis of the maize NAM population, there is evidence that pleiotropic genes contribute resistance to three different fungal foliar blights, but that the MDR phenotype of the NAM founders appears to be largely due to the accumulation of disease specific QTL for each of the three diseases examined. This is expected as the founder lines are elite inbreds

resulting from selection in various breeding programs around the world. As disease pressure varies from region to region, selection for disease resistance has higher or lower priority depending on locality. If a given locality is conducive to disease development, selection for resistance to several different diseases will be a priority and would likely include these three important diseases. In these situations, the breeder has assembled multiple disease resistance phenotypes even though there is limited contribution from a common underlying genetic basis. Contributing to the MDR phenotype, however, are several loci that condition resistance to two of the disease examined. The correlation of allele effects for resistance to two or three diseases across the diverse founders gives evidence for a common underlying gene. Nested association mapping using the MDR index identified three known resistance genes along with numerous other candidate genes and genes of unknown function at the site of SNP associations. Further study of these locations, including dissection of the region in a map-based cloning approach, will be needed to confirm and identify the MDR QTL.

REFERENCES

Balint-Kurti, P. J., J. C. Zwonitzer, M. E. Pe, G. Pea, M. Lee and A. J. Cardinal (2008). Identification of Quantitative Trait Loci for Resistance to Southern Leaf Blight and Days to Anthesis in Two Maize Recombinant Inbred Line Populations. *Phytopathology* **98** (3): 315-320.

Balint-Kurti, P. J., J. C. Zwonitzer, R. J. Wisser, M. L. Carson, M. Oropeza-Rosas, J. B. Holland and S. J. Szalma (2007). Precise Mapping of Quantitative Trait Loci for Resistance to Southern Leaf Blight, Caused by Cochliobolus Heterostrophus Race O, and Flowering Time Using Advanced Intercross Maize Lines. *Genetics* **176**: 645-657.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J. C. Glaubitz, M. M. Goodman, C. Harjes, K. Guill, D. E. Kroon, S. Larsson, N. K. Lepak, H. Li, S. E. Mitchell, G. Pressoir, J. A. Peiffer, M. O. Rosas, T. R. Rocheford, M. C. Romay, S. Romero, S. Salvo, H. S. Villeda, H. Sofia da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich and M. D. McMullen (2009). The Genetic Architecture of Maize Flowering Time. *Science* **325** (5941): 714-718.

Century, K. S., E. B. Holub and B. J. Staskawicz (1995). Ndr1, a Locus of Arabidopsis Thaliana That Is Required for Disease Resistance to Both a Bacterial and a Fungal Pathogen. *Proceedings of the National Academy of Sciences* **92** (14): 6597-6601.

Fokunang, C. N., C. N. Akem, A. G. O. Dixon and T. Ikotun (2000). Evaluation of a Cassava Germplasm Collection for Reaction to Three Major Diseases and the Effect on Yield. *Genetic Resources and Crop Evolution* **47** (1): 63-71.

Fukuoka, S., N. Saka, H. Koga, K. Ono, T. Shimizu, K. Ebana, N. Hayashi, A. Takahashi, H. Hirochika, K. Okuno and M. Yano (2009). Loss of Function of a Proline-Containing Protein Confers Durable Disease Resistance in Rice. *Science* **325** (5943): 998-1001.

Gomez-Gomez, L. and T. Boller (2000). Fls2: An Lrr Receptor-Like Kinase Involved in the Perception of the Bacterial Elicitor Flagellin in Arabidopsis. *Molecular Cell* **5** (6): 1003-1011.

Krattinger, S. G., E. S. Lagudah, W. Spielmeyer, R. P. Singh, J. Huerta-Espino, H. McFadden, E. Bossolini, L. L. Selter and B. Keller (2009). A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. *Science* **323** (5919): 1360-1363.

Lorang, J. M., T. A. Sweat and T. J. Wolpert (2007). Plant Disease Susceptibility Conferred by a "Resistance" Gene. *Proceedings of the National Academy of Sciences, USA* **104** (37): 14861-14866.

Manosalva, P. M., R. M. Davidson, B. Liu, X. Zhu, S. H. Hulbert, H. Leung and J. E. Leach (2009). A Germin-Like Protein Gene Family Functions as a Complex Quantitative Trait Locus Conferring Broad-Spectrum Disease Resistance in Rice. *PLANT PHYSIOLOGY* **149** (1): 286-296.

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S. E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M. O. Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J. C. Glaubitz, M. Goodman, D. Ware, J. B. Holland and E. S. Buckler (2009). Genetic Properties of the Maize Nested Association Mapping Population. *Science* **325** (5941): 737-740.

Mitchell-Olds, T., R. V. James, M. J. Palmer and P. H. Williams (1995).

Genetics of *Brassica Rapa* (Syn. *Campestris*). 2. Multiple Disease Resistance to Three Fungal Pathogens: *Peronospora Parasitica, Albugo Candida* and *Leptosphaeria Maculans. Heredity* **75** (4): 362-369.

Mou, Z., W. Fan and X. Dong (2003). Inducers of Plant Systemic Acquired Resistance Regulate Npr1 Function through Redox Changes. *Cell* **113** (7): 935-944.

Murray, S. L., N. Adams, D. J. Kliebenstein, G. J. Loake and K. J. Denby (2005). A Constitutive PR-1::Luciferase Expression Screen Identifies *Arabidopsis* Mutants with Differential Disease Resistance to Both Biotrophic and Necrotrophic Pathogens. *Molecular Plant Pathology* **6** (1): 31-41.

Nagy, E., T.-C. Lee, W. Ramakrishna, Z. Xu, P. Klein, P. SanMiguel, C.-P. Cheng, J. Li, K. Devos, K. Schertz, L. Dunkle and J. Bennetzen (2007). Fine Mapping of the *Pc* Locus of Sorghum Bicolor, a Gene Controlling the Reaction to a Fungal Pathogen and Its Host-Selective Toxin. *Theoretical and Applied Genetics* **114** (6): 961-970.

Poland, J. A., P. J. Balint-Kurti, R. J. Wisser, R. C. Pratt and R. J. Nelson (2009). Shades of Gray: The World of Quantitative Disease Resistance. *Trends in Plant Science* **14** (1): 21-29.

R Development Core Team (2009). R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Saghai Maroof, M. A., S. W. Van Scoyoc, Y. G. Yu and E. L. Stromberg (1993). Gray Leaf Spot Disease of Maize: Rating Methodology and Inbred Line Evaluation. *Plant Disease* **77**: 583-587.

Salvi, S., G. Sponza, M. Morgante, D. Tomes, X. Niu, K. A. Fengler, R. Meeley, E. V. Ananiev, S. Svitashev, E. Bruggemann, B. Li, C. F. Hainey, S. Radovic, G. Zaina, J. A. Rafalski, S. V. Tingey, G.-H. Miao, R. L. Phillips and R. Tuberosa (2007). Conserved Noncoding Genomic Sequences Associated with a Flowering-Time Quantitative Trait Locus in Maize. *Proceedings of the National Academy of Sciences* **104** (27): 11376-11381.

Scheet, P. and M. Stephens (2006). A Fast and Flexible Statistical Model for Large-Scale Population Genotype Data: Applications to Inferring Missing Genotypes and Haplotypic Phase. **78** (4): 629-644.

Tierens, K. F. M. J., B. P. H. J. Thomma, R. P. Bari, M. Garmier, K. Eggermont, M. Brouwer, I. A. M. A. Penninckx, W. F. Broekaert and B. P. A. Cammue (2002). *Esa1*, an *Arabidopsis* Mutant with Enhanced Susceptibility to a Range of Necrotrophic Fungal Pathogens, Shows a Distorted Induction of Defense Responses by Reactive Oxygen Generating Compounds. *Plant Journal* **29** (2): 131-140.

Wan, J., X.-C. Zhang, D. Neece, K. M. Ramonell, S. Clough, S.-y. Kim, M. G. Stacey and G. Stacey (2008). A *LysM* Receptor-Like Kinase Plays a Critical Role in Chitin Signaling and Fungal Resistance in *Arabidopsis*. *The Plant Cell* **20** (2): 471-481.

Wang, G.-L., D.-L. Ruan, W.-Y. Song, S. Sideris, L. Chen, L.-Y. Pi, S. Zhang, Z. Zhang, C. Fauquet, B. S. Gaut, M. C. Whalen and P. C. Ronald (1998). *Xa21d* Encodes a Receptor-Like Molecule with a Leucine-Rich Repeat Domain That Determines Race-Specific Recognition and Is Subject to Adaptive Evolution. *THE PLANT CELL* **10** (5): 765-780.

Wisser, R. J., P. J. Balint-Kurti and R. J. Nelson (2006). The Genetic

Architecture of Disease Resistance in Maize: A Synthesis of Published Studies. *Phytopathology* **96** (2): 120-129.

Wisser, R. J., Q. Sun, S. H. Hulbert, S. Kresovich and R. J. Nelson (2005). Identification and Characterization of Regions of the Rice Genome Associated with Broad-Spectrum, Quantitative Disease Resistance. *Genetics* **169** (4): 2277-2293.

CHAPTER 4

In the eye of the beholder: The effect of scorer variability and different rating scales on QTL mapping and genomic selection

ABSTRACT

The agronomic importance of developing durably resistant cultivars has led to considerable research in the field of quantitative disease resistance (QDR) and, in particular, mapping quantitative resistance loci (QTL). The assessment of QDR is typically conducted by visual estimation of disease severity, which raises concern over the accuracy and precision of visual estimates. While previous studies have examined the factors affecting the accuracy and precision of visual disease assessment in relation to the true value of disease severity, the impact of this variability on the identification of disease resistance QTL has not been assessed. In this study, the effects of scorer variability and rating scales on mapping QTL for northern leaf blight resistance in maize were evaluated in a recombinant inbred line population. As expected, more experienced individuals had higher precision while the average of multiple individuals gave the highest precision. Using a direct estimation of diseased leaf area produced higher precision than using an ordinal scale. Stepwise general linear model selection (GLM) and inclusive composite interval mapping (ICIM) were used for QTL mapping. For GLM the same QTL were largely found across individuals, though some QTL were only identified by some individuals. Strikingly, the magnitudes of estimated allele effects at identified QTL were drastically different, sometimes by as much as

three fold. ICIM produced highly consistent results across individuals and for the different rating scales in identifying the location of QTL. Genomic selection models were also evaluated for predictive power. We did not find significant differences between the prediction accuracy of different scales though there were differences between individuals. We conclude that the heritability (precision) of quantitative disease resistance can be increased through improved accuracy in visual evaluation, which can be accomplished though increased experience or averaging the scores of multiple individuals. Despite variability between individuals, the identification of QTL was largely consistent between individuals, particularly for ICIM. However, care should be taken in estimating the QTL allele effects as this was highly variable and depended on the individual who conducted the rating. Application of genomic selection models in plant breeding can utilize different rating scales with equal predictive power, though increased precision in individual ratings will lead to better predictions.

INTRODUCTION

Beauty is in the eye of the beholder, as the old saying goes. One would hope, however, that for visual evaluation of quantitative disease resistance (QDR), the results and inferences would not be specific to the individual who did the disease rating or the scale that was used for rating. Due to the importance of QDR, hundreds of studies have been published on the mapping of quantitative trait loci (QTL) for disease resistance in plants (e.g. Wang *et al.* 1994; Poland *et al.* 2009). The populations used in these studies have been evaluated almost exclusively using visual estimation of disease severity. Typically, populations were evaluated by one or a few individuals using several replications repeated over different years, seasons and/or environments (e.g. Baumgarten *et al.* 2007; Balint-Kurti *et al.* 2008). While the results of those studies are only as accurate as the methods used to assess resistance, little is known about how scorer variability and different rating scales influence QTL mapping and genomic selection.

QDR is an important objective in crop breeding programs as this type of resistance tends to be more durable (Poland *et al.* 2009). For some pathosystems, QDR is the only available type of resistance. The development of durably resistant cultivars that carry QDR remains challenging due to the polygenic nature of the phenotype and the small phenotypic effects of individual genes.

Progress in resistance breeding requires accurate methods for rating QDR. In breeding programs, selection for resistance could be achieved by focusing on the most resistant end of the spectrum. When truncated selection is practiced, as in most breeding programs, it is not important to accurately

rate the more susceptible material. With the implementation of genomic selection models, in contrast, accurately evaluating the whole spectrum of the germplasm will have increased importance in order to improve the prediction accuracy of the model. In the field of quantitative genetics, accurate ratings have importance in providing precise measures of QTL positions and effects. Increased accuracy and precision will lead to improved heritability. This results in improved selection efficiency and more power in detecting QTL.

The accuracy and precision of visual evaluation of disease levels have been analyzed in numerous previous studies (e.g. O'Brien and van Bruggen 1992; Nutter *et al.* 1993; Nita *et al.* 2003; Hartung and Piepho 2007; Bock *et al.* 2008). Accuracy is generally defined as the closeness of the visual estimate to the true level of disease. Precision is a measure of the consistency of visual evaluations (Nutter *et al.* 2006). **Figure 4.1** gives an example of three ratings with different levels of accuracy and precision. Many previous studies have focused on evaluation of computer simulated diseased leaves where the true underlying percentage of diseased area is known (e.g.Hartung and Piepho 2007) or evaluation of single leaves where image analysis was used to determine the true percentage (e.g.Bock *et al.* 2008). While rating of simulated diseased leaves allows precise interpretation because the real underlying percentage of disease level is known, there may be little relationship to actual field evaluation of disease levels, for which entire plants and whole plots must be taken into account.

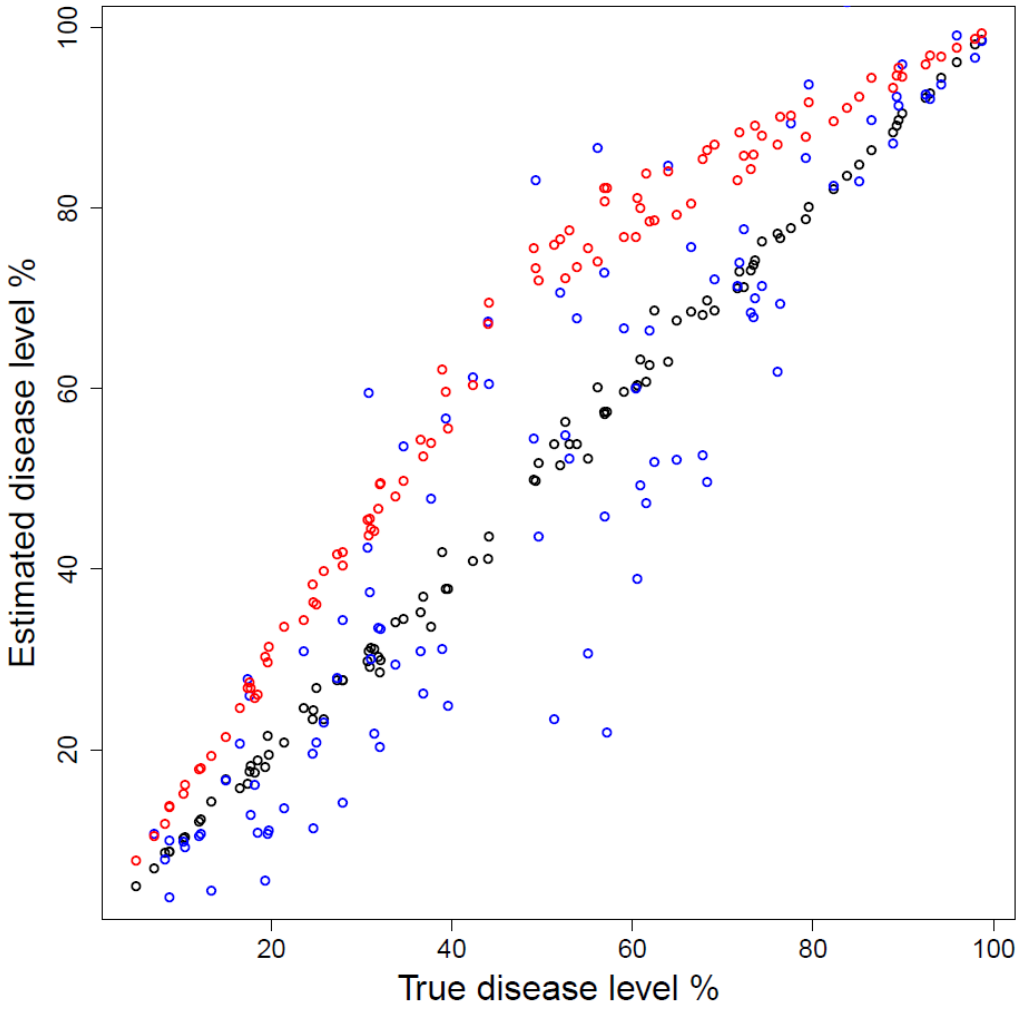


Figure. 4.1. Example of accuracy and precision in disease severity estimates.

Estimates were simulated for three different “hypothetical scorers” and are plotted against the true value of disease severity. “Scorer 1” shows both high accuracy and precision (black). “Scorer 2” (blue) shows high accuracy (estimates are centered around true value) but lack of precision, while “Scorer 3” has high precision (estimates are highly repeatable) but low accuracy.

Many inferences have been made through previous studies examining the concepts of visual ratings. By comparing visual evaluation of diseased leaves with image analysis of the same leaves, Parker et al (1995) found that levels of *Septoria tritici* leaf blotch and powdery mildew were not accurately or consistently estimated by several experienced individuals (Parker *et al.* 1995). Over-estimation was evident at low disease levels. The correlation between estimates of yield loss and disease severity for corky root in lettuce was found to vary with different qualitative and quantitative scales (O'Brien and van Bruggen 1992). The qualitative scales were found to be most precise, while the quantitative scale correlated best with yield loss. For *Phomopsis* leaf blight of strawberry, the correlation between individual estimations and the actual disease severity was high for six different individuals using either the H-B scale (0.85 to 1.00) or direct percentage estimation (0.92 to 0.99), though the direct percentage estimation was more accurate (Nita *et al.* 2003). Estimations were centered on the actual value, indicating a lack of precision but no bias in visual evaluation.

Horsfall-Barratt (1945) originally proposed a modified ordinal disease rating scale that was based on logarithmic increases in disease severity. Each unit increase in the ordinal scale represented a doubling in the disease severity. This scale was proposed because it was believed that the reference stimulus (i.e. amount of disease) needed to double for an individual to be able to detect a visual difference. It is now acknowledged, however, that linear increases in visual stimulus can be detected (Nutter and Esker 2006). This concept was examined in the grape powdery mildew and wheat leaf rust pathosystems, and is presumed to hold true for other pathosystems.

The concept of '**J**ust **N**oticeable **D**ifference' or JND was developed in the field of psychophysics to describe the minimum difference that can be perceived between a reference stimulus and a comparison stimulus (Nutter and Esker 2006). The JND was proposed by E. H. Weber and states that *JND = kS*, where Weber's constant *k* depends on the type of stimulus and S is the magnitude of the stimulus (Nutter and Esker 2006). In other words, the minimum difference that can be reliably detected between two different stimuli (i.e. diseased leaf) is proportional to the magnitude of the stimulus (i.e. diseased leaf area). This concept has been examined in the field of plant pathology and shown to hold true for distinguishing differences between diseased leaf samples (Nutter and Esker 2006). It was found that Weber's fraction was in the range of 0.16 and 0.21 for leaf rust in wheat and 0.11 to 0.26 for downy mildew of grape. This indicates that for visual evaluations of disease severity, individuals were generally able to distinguish a relative difference of about 20% (e.g. the relative difference between 20% vs. 24% and between 50% vs. 60% is 20% in both cases, but the absolute differences are 4% and 10% respectively). Therefore, it is expected that for actual disease severity in the range of 40-60%, visual estimates will be less precise than for lower disease levels of 5-20%.

Experience and training have been shown to improve the accuracy of disease assessment (Bock *et al.* 2009). Nutter *et al.* (2006) showed that computer training programs that simulated diseased leaves can increase the accuracy and the precision of disease assessment ratings. Additionally, the use of a standard area diagram can improve the accuracy of visual disease ratings (Nutter *et al.* 2006). Depending on the pathosystem under study, this approach has been used to help improve accuracy and precision between

individuals.

Hartung and Piepho (2007) conducted a simulation-based study comparing the accuracy, precision, and time needed between direct percentage ratings and ordinal scales. More accurate results were obtained using a percentage scale than using a 1-9 ordinal scale for both trained and untrained individuals. Use of an ordinal scale was faster for untrained individuals while a percentage scale based on 5% increments was fastest for trained individuals. They also described the psychological impediments to using a percentage rating scale. Although most individuals indicated that giving a direct evaluation of disease percentage seemed less accurate, the percentage rating produced more accurate results.

Image analysis and remote sensing have been proposed and utilized as tools for accurately determining disease levels. Image analysis has been used to identify QTL for Gibberella stalk rot resistance, though the results were not compared to visual evaluation of the trait (Pè *et al.* 1993). Image analysis has been used to analyze and map QTL for other traits such as kernel morphology (Campbell *et al.* 1999) and flour-milling yield (Berman *et al.* 1996). While advances have been made in utilizing both image analysis and remote sensing, the constraints of typical breeding programs still limit the utility of both approaches. Breeding programs typically test thousands of entries in early-generation trials. These trials are often conducted in small plots that are not sufficient in area for remote sensing (or remote image analysis). Concurrently, the large number of lines that need to be evaluated limit the time available for evaluation of each line. This constrains the utility of image analysis (based on current technology), which requires careful positioning or collection of leaf samples prior to taking an image. While an individual can rate a disease plot

in 10-20 seconds, the additional time needed to collect good images for analysis (say a speedy rate of 5 minutes per line) would reduce the number of evaluations over a day (6 hours) from around 1800 to less than 100. This is a considerable impediment to evaluating a large number of lines. Thus, for the time being, visual rating remains the most practical approach to disease assessment in most breeding and genetics programs.

While multiple studies have looked at the ability of individuals to give accurate and precise visual assessments of disease as well as differences between visual rating scales, there has been limited work looking at the impact of variability in visual disease estimation on the results and inferences made from such studies. In particular, there have been many reports of mapping quantitative resistance loci (QTL) in plants. Usually, the objective is to identify genes underlying this complex trait and/or tagging QTL with molecular markers for selection in a breeding program. To examine impact of scorer variability on the results from these studies, we examined the genomic position and relative effect of QTL identified by different scorers. To do so, we utilized the maize – *Setosphaeria turcica* (anamorph *Exserohilum turcicum*) pathosystem. *S. turcica* is the causal agent of northern leaf blight (NLB), an economically important maize disease throughout the world. The pathogen spreads through the plant vasculature, causing large necrotic lesions on the leaves. Under the maize diversity project, excellent genetic resources have been developed for the maize community, providing tools for study of quantitative traits in maize. The combination of well designed maize populations, the importance of NLB, and the quantitative genetic expertise found in the maize community makes this pathosystem an excellent model for studying quantitative disease resistance in plants.

MATERIALS AND METHODS

Plant Materials. The maize nested association mapping population (NAM) is a set of 25 recombinant inbred line (RIL) families that were derived by crossing each of 25 diverse inbred lines with a common reference inbred line (Yu *et al.* 2008; Buckler *et al.* 2009; McMullen *et al.* 2009). From the 25 RIL families, the MS71 x B73 population was previously identified as having large variation in quantitative resistance to NLB, and minimal variation for relative maturity (Poland *et al.* in prep.). Minimal variation in relative maturity is beneficial as quantitative disease resistance in plants has been previously associated with relative maturity. Seed for the MS71 x B73 RIL population was generously supplied by E. Bucker. The population consists of 200 S_5 RILs, of which 191 were used for this study. The two inbred parents, MS71 and B73 were used as checks throughout the experiments. The population is genotyped with 1106 SNP markers of which 701 are polymorphic and used for mapping. Marker positions based on the NAM composite map were used (McMullen *et al.* 2009).

Scorers. Twenty two persons volunteered to participate in the scoring experiment in 2008, 2009 or both years. Scorers included undergraduates, graduate students, and faculty, in Plant Science, Plant Breeding and/or Plant Pathology. Though not all had experience in plant disease rating, all had some experience in plant biology related research. To gauge the amount of experience in rating plant disease, an informal questionnaire with three questions was presented to each scorer asking about previous experience. Scorers were asked to rate themselves as follows; 1) no experience 2) little

experience 3) some experience 4) experienced 5) very experienced in two areas: a) experience scoring plant disease in general and b) experience scoring northern leaf blight.

Field Trials for Northern Leaf Blight. Field trials were conducted during 2008 and 2009 at the Robert B. Musgrave Research Farm in Aurora, New York. Trials were planted on May 14, 2008 and May 18, 2009. Lines were planted as single row plots 7' in length with 30" between rows. Plots were overplanted and thinned to ~10 plants/row. Pre-emergence and post-emergence herbicide applications were applied each year. Trials were laid out in an augmented incomplete block design with one replication in 2007 and two replications in 2008. Each block consisted of 20 RILs and two checks (B73 and MS71). Each year the trials were artificially inoculated with *S. turcica* Race 1 (isolate NY-001 from the lab of R. Nelson). Plants were inoculated at about the 6 to 8 leaf stage, which corresponded to June 27, 2008 and July 16, 2009. The 2009 season had an extremely cool spring and plant growth was very delayed compared with 2008. Each year, two types of inoculums were simultaneously applied 1) 2.5–3.0 ml of dried infected sorghum grains, previously inoculated and cultured for 2 weeks, and 2) 1.0 ml of a spore suspension (1×10^3 conidia/ml) in H_2O with 0.02% Tween 20. For the spore suspension, *S. turcica* isolate NY001 was cultured on lactose casein agar (LCA) plates for 2-3 weeks at room temperature under 12 hr light / 12 hr dark conditions. Spores were harvested by flooding the plates with sterile H_20 and scraping with a glass rod and filtering through cheese cloth. Spore concentrations were determined and diluted to 1×10^3 conidia/ml.

Phenotypic Evaluation. Trials were visually evaluated for disease severity at two time points each year (ratings 1 and 2). Disease severity was

assessed as diseased leaf area defined as the percentage of total leaf area in the plot that was covered by necrotic lesions from NLB. There was limited senescence at the time of rating, though individuals were advised to take that into account when rating. There was no significant presence of secondary diseases to complicate scoring. In 2009 two different rating scales were used, a direct percentage scale (%) and a 0-9 logarithmic based scale (0-9). In 2008 only the % scale was used. For the percentage rating, disease severity was estimated as the percentage of total leaf area necrotic with disease using a 0-100% scale with 1% increments. For the 0-9 scale, a typical disease severity scale based on logarithmic increases in diseased leaf area was used (Hartung and Piepho 2007) and additional semi-qualitative descriptors were incorporated (**Table 4.1**)

Table 4.1. Ordinal rating scale (0-9) used for rating northern leaf blight disease severity.

The ordered classes were based on a slightly modified logarithmic scale. Additional descriptions of the expected phenotype for each class were added to aid in scoring and consistency.

Northern Leaf Blight 0-9 scale

Category	DLA percentage range	Additional Descriptors
0	0%	no lesions visible
1	<1%	few small lesions on lower leaves
2	1-2%	several lesions on lower leaves
3	3-5%	many lesions on lower leaves
4	5-8%	coalescent lesions on lower leaves
5	8-12%	lower leaves mostly blighted, few small lesions on middle leaves
6	12-20%	lower leaves almost completely blighted, some lesions on middle leaves
7	20-33%	lower leaves completely blighted, considerable lesions on middle leaves
8	33-66%	lower leaves completely blighted, middle and ear leaf largely blighted
9	66-100%	most to all of green leaf tissue blighted

In 2008, four scorers participated in the first rating and six scorers in both ratings. In 2008, all scorers used a percentage rating scale. In 2009, 12 scorers evaluated the population at both ratings and four scorers for only the second. In 2009, four scorers rated using both the % and the 0-9 scale. The individuals, ratings, and scale used are listed in **Table 4.2**.

Data Analysis. In this study we consider the correlation between replications and the correlation between different individuals and the mean rating as a measure of precision. For accuracy it will not be possible to determine the true level of disease so any inferences about accuracy will be difficult to make.

Correlations between individuals and between replications for individuals were evaluated in JMP v7.0 using the multivariate methods analysis. Statistical tests based on correlation values were done in JMP by conducting t-test or fitting a general linear model where noted.

Best linear unbiased predictions (BLUPs) for each line were determined using PROC MIXED in SAS v9.1. A single model was fit for each of the rating scales which incorporated all of the ratings from all scorers. The full mixed model was:

$$Y_r = S_n + Y_i + R_{j(i)} + B_{k(j)} + L_m$$

where Y_r is the multivariate response of disease severity value at rating *r*, S_n is the random effect of scorer *n*, Y_i is the random effect of year *i*, $R_{j(i)}$ is the random effect of replication *j* within Year *i*, $B_{k(j)}$ is the random effect of block *k* in replication *j* and L_m is the random effect of line *m*. The model solution gave BLUPs for each line.

Table 4.2. Scores , scale and ratings

The rating scale used, the evaluations conducted and the disease rating experience of the 22 individuals who participated in the study. There were two rating time points during year 1 (2008) and year 2 (2009). Experience was self-assessed using a 1-5 scale with 1 representing no experience and 5 being very experienced rating plant diseases in general and maize northern leaf light (NLB) respectively.

	trait_id	Percentage				1-9 scale		Experience	
	year	2008		2009		2009			
	rating	1	2	1	2	1	2	General	Maize NLB
Individual	1					x	x	4	1
	2			x	x	x	x	1	1
	3			x	x	x	x	1	1
	4	x	x	x	x			3	5
	5			x	x	x	x	2	1
	6	x						1	1
	7	x	x	x	x	x	x	3	5
	8	x		x	x			4	5
	9	x	x					1	1
	10	x	x					1	1
	11			x	x			2	3
	12	x						1	1
	13	x	x					2	3
	14				x			4	1
	15						x	4	1
	16	x						5	3
	17					x	x	2	2
	18	x	x	x	x			5	2
	19				x			3	1
	20						x	3	1
	21			x	x			2	1
	22	x						1	1

For reduced models where certain individuals, years, ratings, or replications were to be analyzed, the random effect terms were left out accordingly. For example, the mixed model for Scorer 1 at Rating 2 in Year 2 would be as follows:

$$Y = R_j + B_{k(j)} + L_m$$

where Y has become the univariate response of disease severity estimation values at Rating 2 in Year 2, R_j is the random effect of replication *j*, $B_{k(j)}$ is the random effect of block *k* in replication *j* and L_m is the random effect of line *m*.

Quantitative Trait Loci Mapping. Two methods for QTL mapping were employed. Stepwise General Linear Model (GLM) selection was conducted in SAS v9.1.3 (PROC GLMSelect). The BLUPs from each respective scorer, year, rating combination (and the BLUPs from the full model) were assigned as the response variables and model selection was conducted for marker effects. Marker genotypes were assigned as 0, 1, and 2 for B73, heterozygous and MS71 genotypes, respectively. Marker effects were fit as a continuous variable consistent with an additive effects model. A selection threshold of p-value = 0.001 was used for entry and removal of selected effect in the model. Model solutions were saved for the estimated effects of each selected marker in the model.

Inclusive composite interval mapping (ICIM) (Li *et al.* 2007) was conducted using QGene v.4.2.3 (Joehanes and Nelson 2008). Cofactor selection was conducted with stepwise selection with a selection threshold of F = 3.0. The default value of 2cM was used for a scan interval. As with the GLM, the BLUPs from each of the respective scorers were mapped as a separate trait. Due to computational constraints, permutation analysis to determine an experimental significance threshold for each individual and rating

was not conducted. An arbitrary threshold of LOD=3 was used as a general threshold for significant QTLs.

Genomic Selection. Each respective set of phenotype evaluations were evaluated using a ridge regression genomic selection model (Meuwissen *et al.* 2001). The ridge regression model was programmed in Java (Hiroyoshi Iwata, unpublished). In each case, cross-validation was conducted by iteratively removing one line from the prediction model and then evaluating the correlation between the predicted effect for the single line omitted from the dataset and the empirical phenotypic data observation.

RESULTS

Accuracy and precision of disease ratings. We examined all possible pair-wise correlations for each of the respective years. Correlations were generally high and all pair-wise correlations examined were significant at p-value < 0.0001. Though the % scale used continuous values from 1 to 100, there was a tendency to score using intervals of 5 (**FIG. 4.2**). This was particularly evident at higher disease levels.

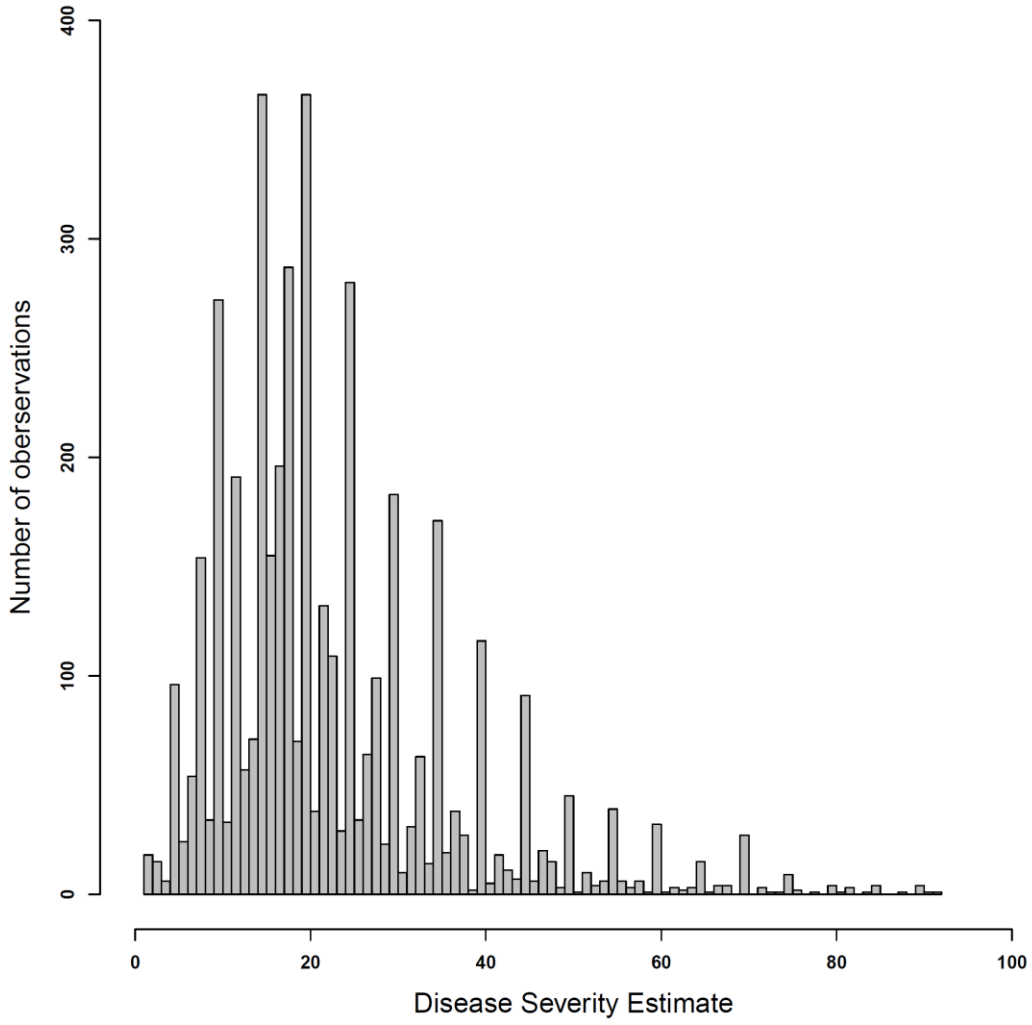


Figure 4.2. Histogram showing tendency to score in intervals of 5 when using the % rating scale.

Histogram showing tendency to score using intervals of 5. The number of disease severity estimations for each percentage point of 1-100 are shown. This tendency has been observed in previous studies on the accuracy of visual evaluation of disease resistance. Scoring at intervals of 5 creates pseudo-bins or a type of ordinal scale.

To examine the precision of different individuals, the correlations between replications for the Year 2 trial were examined (**Figure. 4.3**). The correlations varied from 0.409 to 0.871 showing considerable differences between individuals. It was observed that there were generally higher correlations between replications for individuals when using the % than the 0-9 scale. The average correlation when using the % scale was 0.764 while the mean correlation for individuals using the 0-9 scale was only 0.603 (**Figure. 4.3**). A multivariate linear model was fit to the replication by replication correlations to test the effect of rating scale, individual, experience, and rating time-point. There was a significant effect of rating scale (p-value < 0.0001) with the % scale being more precise than the 0-9 scale. More experience with scoring NLB resulted in increased precision (p-value = 0.0009), while there was only a trend for general disease scoring experience (p-value = 0.0967). There was also an increase in precision for the second disease rating (p-value = 0.001).

To further examine the effect of experience on consistency, the correlations between individuals were compared to the experience of the two respective individuals. The pair-wise correlation for a given pair was modeled against the sum of the experience scores for the two respective individuals (**Figure 4.4**). A general linear model was fit for the pair-wise correlation against the sum of the respective scores for the two individuals accounting for rating scale as a covariate. Again there was a significant trend for each measure of experience (general, NLB, and combined) with each being highly significant (p-value < 0.0001).

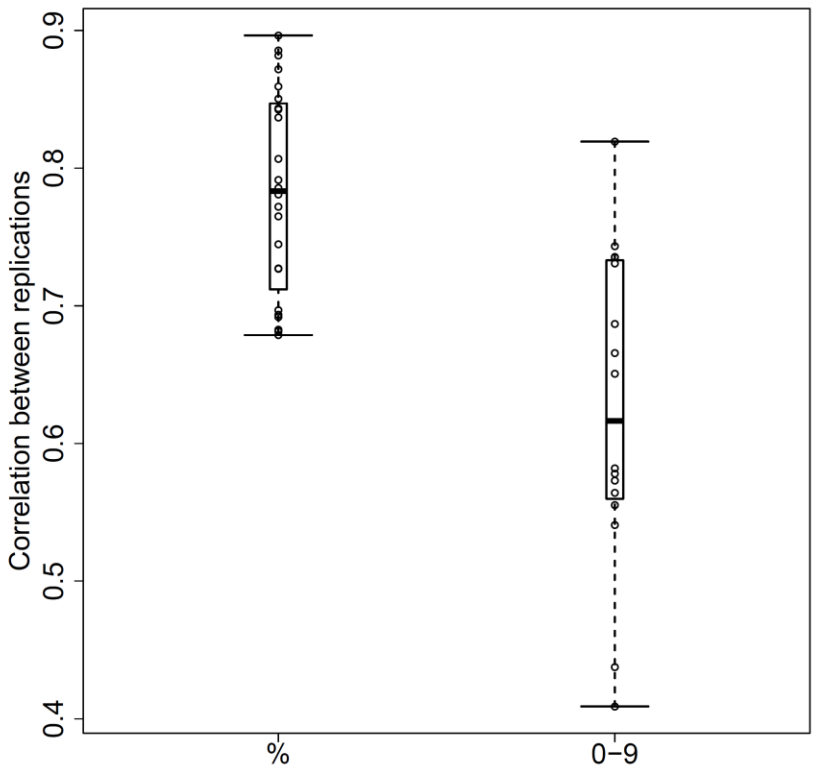


Figure. 4.3 Higher accuracy results from using % scale than 0-9 ordinal scale

Comparision of accuracy of % and 0-9 scales based on within-scorer correlation between replications in Year 2. Individual correlation values are shown as circles with the box plot showing the 25th and 75th percentiles. The difference between the means of the two rating scales is significant (p-value < 0.0001).

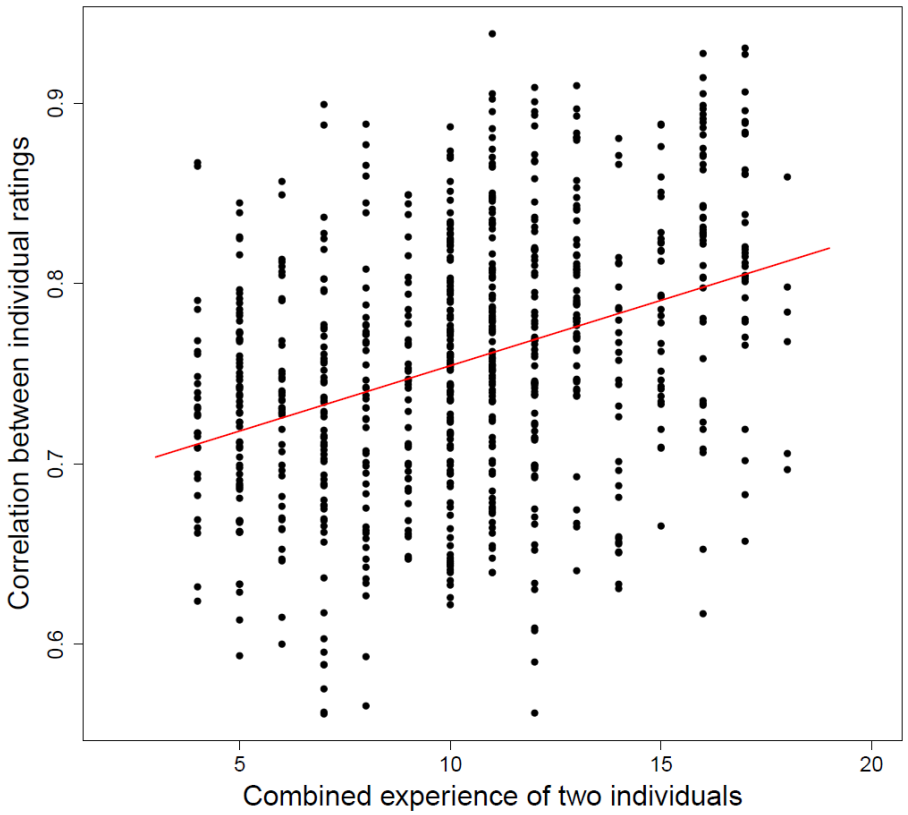


Figure 4.4. Effect of experience on consistency between two different individuals for the % scale

Each plotted point represents the pair-wise correlation between the ratings of two different individuals. The correlation between the two individuals is plotted against the combined experience of the two individuals. The combined experience is presented as the sum of experience values (1-5) of each individual for both general disease rating experience and maize NLB rating experience (two experience values for each individual with minimum combined value of 4 and maximum value of 20). The trend is significant (p-value < 0.0001) with $R^2 = 0.127$. The same trend was evident for the 0-9 scale.

With disease progress during the time between the two ratings the correlation between ratings should not be expected to be complete. However, with disease progress and since lines were evaluated after flowering, is it not expected that any ratings at the second time point would be lower than the previous scoring. For scorers who conducted two ratings, the number of lines that were rated lower for the second rating was determined. There were large differences with some scorers rating fewer than 1% of the lines lower on the second rating while other scorers rated over 50% of the lines lower at the second rating. This is another indicator of precision as a lower score for the second rating means either the first rating was overestimated, the second rating was underestimated, or both.

To examine the congruency of the direct percentage scale and the ordinal scale (with underlying percentage ranges), the % and 0-9 scores for the four individuals who evaluated with both scales were compared. Each class of the 0-9 scale was defined by an underlying percentage range. This range was used to determine if the % ratings assigned to a given plot fell within the defined ranges of the 0-9 scale. The fraction of % ratings outside the defined class ranges varied from 36% to 97%. For the first rating, three of the four individuals had over 70% of the % observations outside of the 0-9 classes. This general lack of congruency between the two scales indicates that comparison of results from different scales is particularly uncertain, even if an ordinal scale has a defined underlying percentage basis. While the correlation between % and 0-9 ratings was strong, there was a lack of direct correspondence.

We observed that visual estimates of NLB disease severity were heteroscedastic; that is, the variability of estimates increased with increasing

disease severity (**Figure 4.5**). This is consistent with previous studies on visual accuracy in disease rating (Hartung and Piepho 2007). For example, with the second rating in Year 2 using the % scale, for a plot that had an average of 10.1, the estimated values ranged from 4 to 20. For a second plot that averaged 49.9, the estimated values ranged from 20 to 70. This same trend was not apparent in the 0-9 ratings (**Figure 4.6**). The increasing variability was easily seen in a residual plot of a linear fit of individual % estimations on the average of all estimations. Again, the uniform nature of the 0-9 residuals based on the same fit indicates that heteroscedasticity is not a problem with the 0-9 scales.

Consistency of QTL identification between individuals and rating scales. To identify molecular markers associated with NLB resistance, stepwise general linear model selection was conducted for each of the respective scorer, year, rating, and scale combinations. These results show a general consistency for model selection of the same QTL. Several QTL were identified across all scorers. However, there was discrepancy between individuals on the identification of smaller effect QTL (**Figure 4.7**). To best compare the difference between rating scales, the BLUPs from all individuals were used for mapping with the %, 0-9 and also a square root transformation of the % scale (**Figure 4.8**). As the % and 0-9 rating scales have a different variance the standardized allele effects were used to compare the results from these two scales. There were 11 QTL identified for the 0-9 and the sqrt% scale while 9 QTL were identified for the % scale. Eight of these QTL were identified across all three scales and the standardized allele effect estimates were roughly equivalent. Two QTL were identified in both the 0-9 and sqrt% scales indicating the detection of smaller effect QTL might be sensitive to the

type of scale and the resulting distribution. The additional two QTL were found in only one rating scale and might also be false positives. The same trend was seen when ratings from different scorers were used for mapping. Several of the larger effect QTL were identified across all scorers, while smaller effect QTL were identified only by some (**Figure 4.8**).

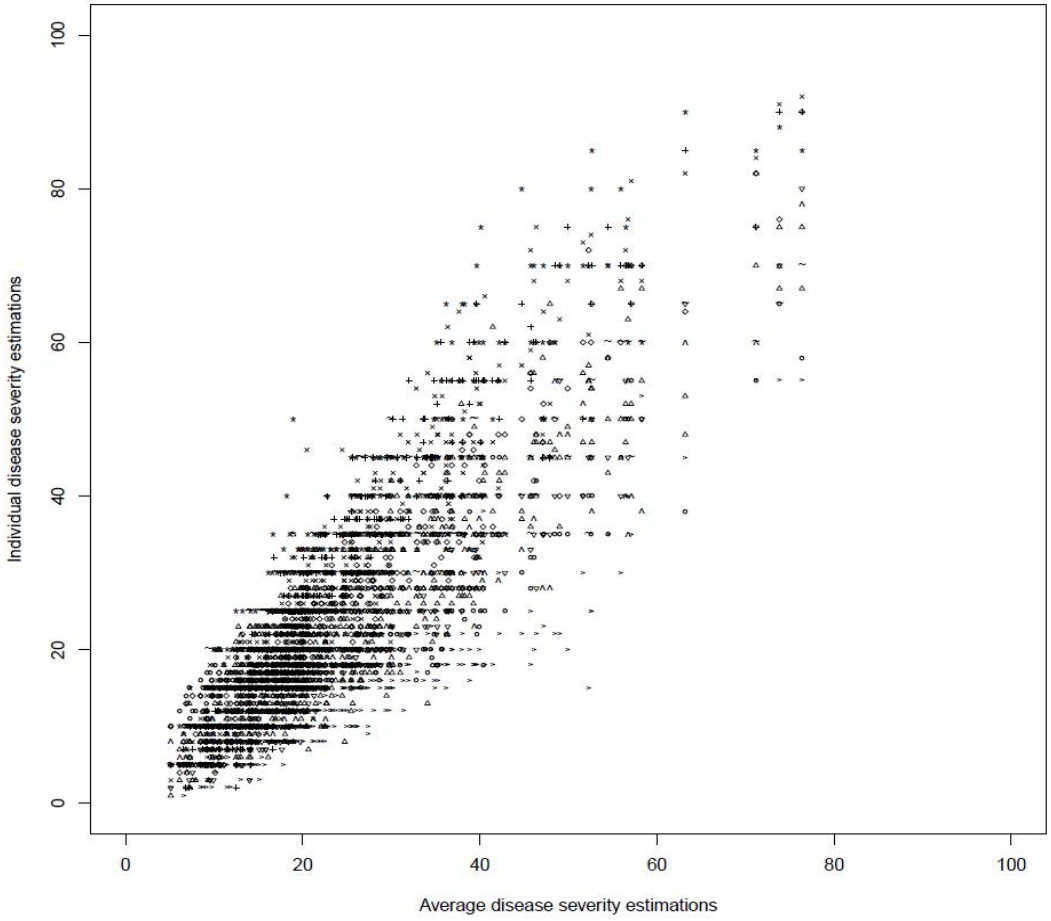


Figure 4.5 Accuracy of disease severity ratings decrease with increased disease severity.

The % rating scale showing increasing variability with increasing disease severity. All individual disease severity estimates plotted against the average of the estimates. Data is from Rating 2 in Year 2 for individuals using the % scale.

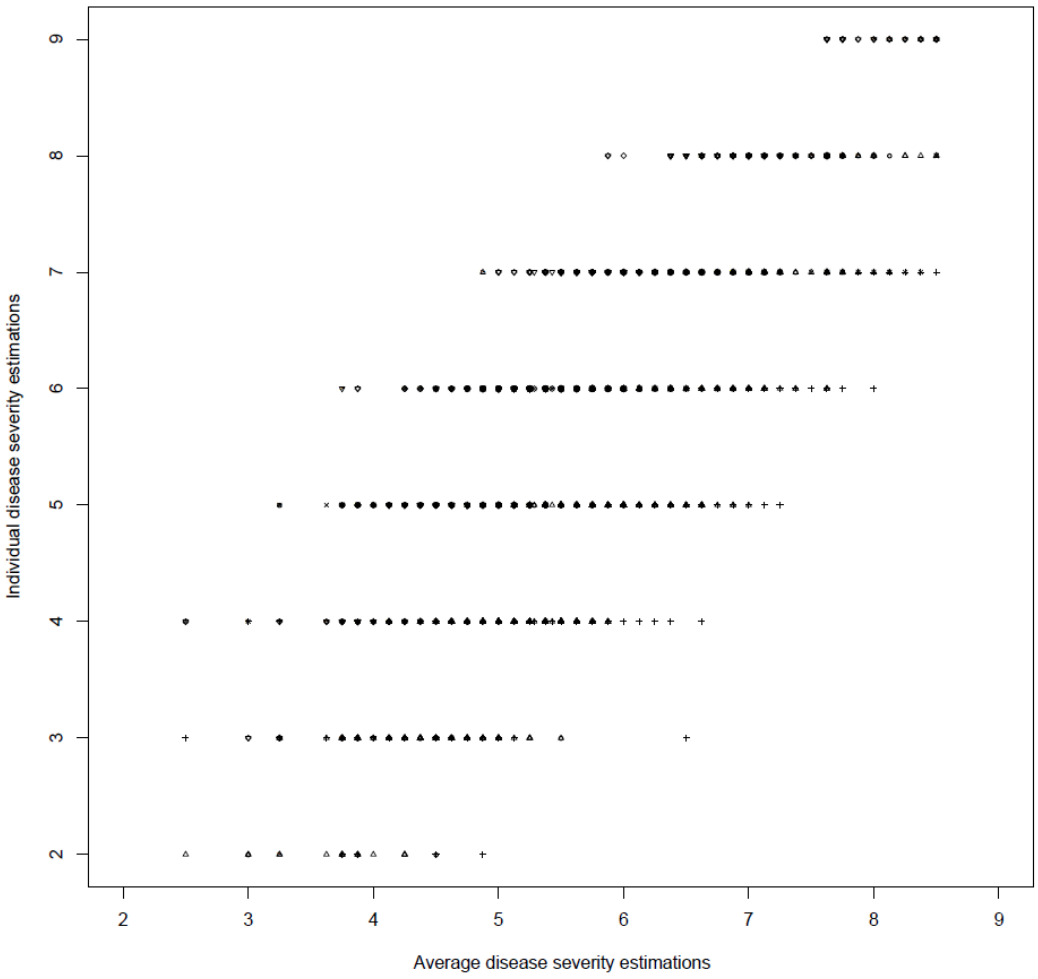


Figure 4.6. Using 0-9 scale leads to uniform error for varying levels of disease severity

The 0-9 scale does not show increasing variability with increasing disease severity. All individual disease severity estimates plotted against the average of the estimates. Data is from Rating 2 in Year 2 for individuals using the % scale.

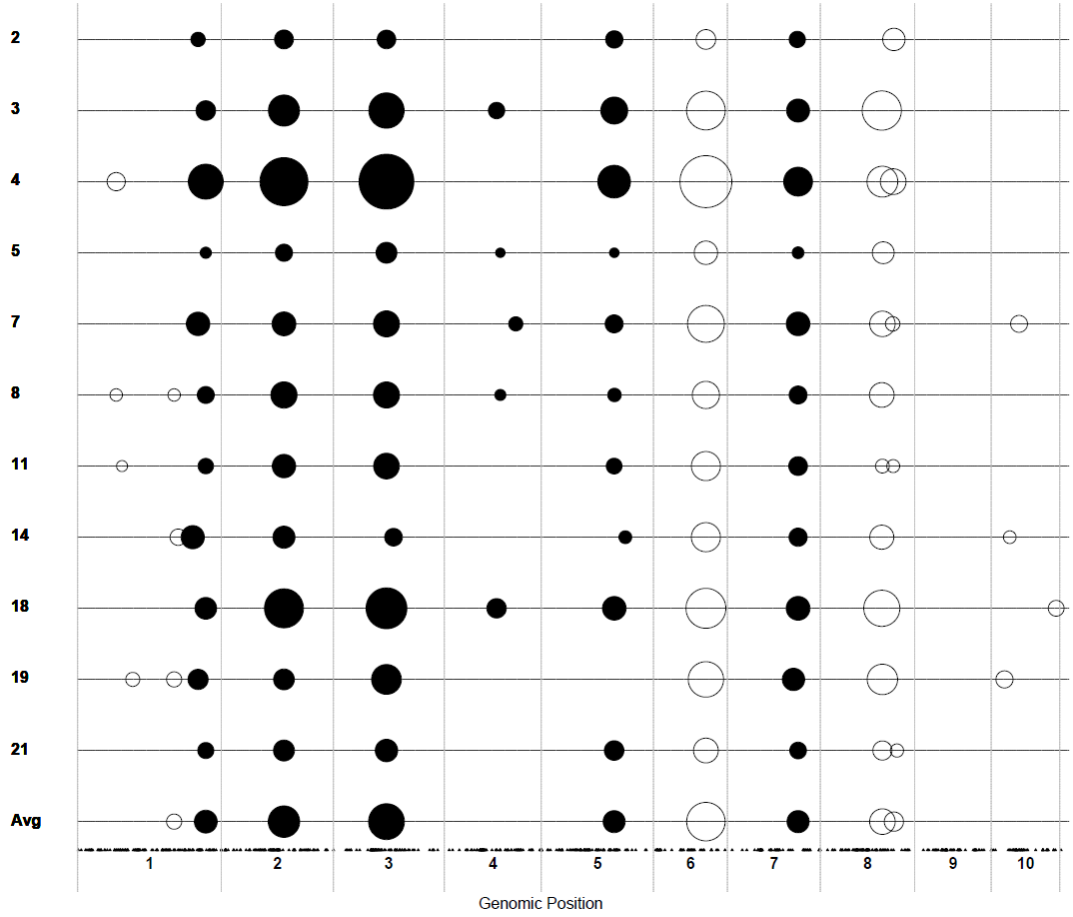


Figure 4.7. Comparison of QTL mapping results between scorers using the % scale

Genomic position of QTL identified from stepwise general linear model selection for each scorer using the % scale. The genomic positions of QTL identified are shown along the horizontal lines with vertical lines showing chromosome breaks. The effect size of identified QTL is represented by the size of the circle. Solid circles represent resistance from Ms71, while open circles represent resistance from B73. The position of molecular markers is shown by black triangles along the bottom.

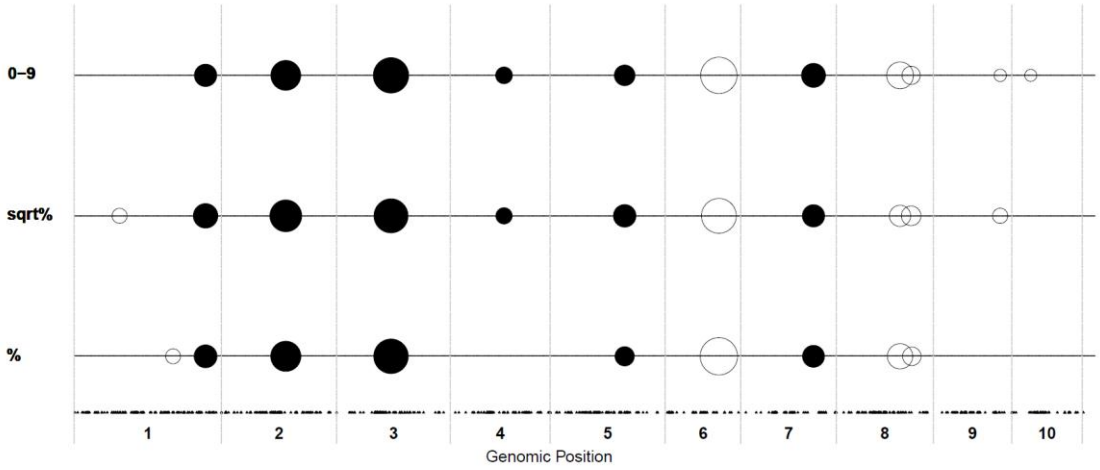


Figure 4.8 Comparison of QTL mapping results for % and 0-9 ordinal scales

Genomic position of QTL identified from stepwise general linear model selection for the overall average using the % and 0-9 rating scales. A sqrt transformation of the % data is also included (sqrt%). The genomic positions of QTL identified are shown along the horizontal lines with vertical lines showing chromosome breaks. The effect size of identified QTL is represented by the size of the circle. Solid circles represent resistance from Ms71, while open circles represent resistance from B73. The position of molecular markers is shown by black triangles along the bottom.

The GLM solutions gave estimates of selected QTL effects as well as standard errors for those estimates. These effect estimates are analogous to the allele effect at that locus. The estimated effects of different individuals for the % scale in Year 2 were compared and showed significant differences for each of the loci that were identified by multiple individuals (**Figure 4.9**). To determine if this was an effect of different variance for the trait distribution between scorers, standardized estimates were compared. These values were much more consistent between scorers though significant differences remained (**Figure 4.10**).

Inclusive composite interval mapping produced very consistent results across different rating scales. As the BLUP of all individuals tended to be the most accurate measure of actual disease levels, these values were also used for comparison of the %, sqrt%, and the 0-9 scales. ICIM using the scales produced almost identical results in terms of QTL position and significance of identified QTL effects (**Figure 4.11**). This is despite a non-linear trend between the % and the 0-9 scale. For ICIM across all individuals, the positions of significant QTL were largely consistent, though the height of the peak varied indicating differences in the power of QTL detection among individuals (**Figure 4.12**). As would be expected, phenotypic observations with lower precision resulted in lower power of QTL detection.

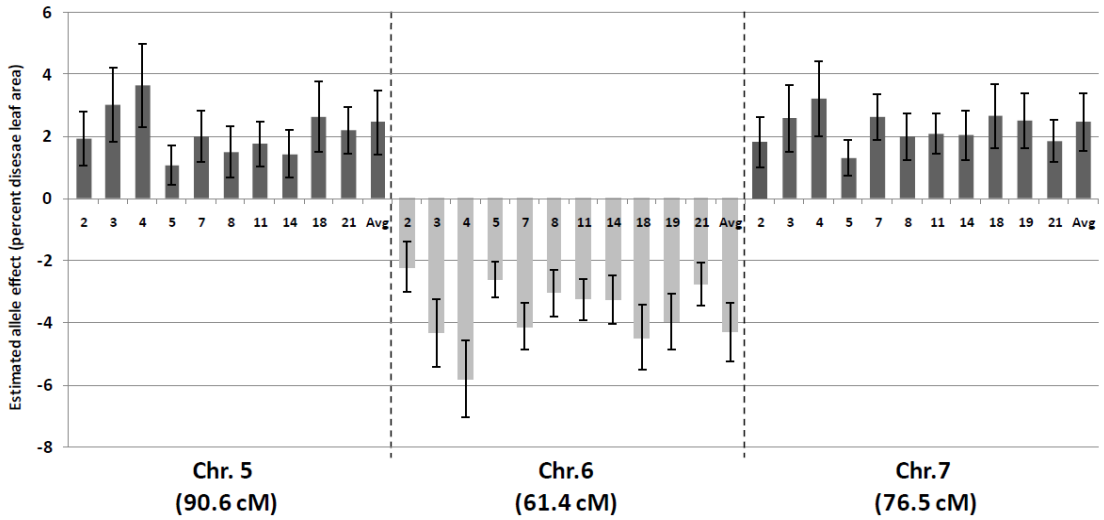


Figure 4.9 Variability of estimated allele effects between scorers

Estimated allele effects for scorers using the % scale at three QTL identified by all individuals. The height of the bar represents the expected percentage increase or decrease in disease conditioned by the MS71 allele based on the ratings for each individual and the average of all individuals. 95% confidence intervals are shown in brackets.

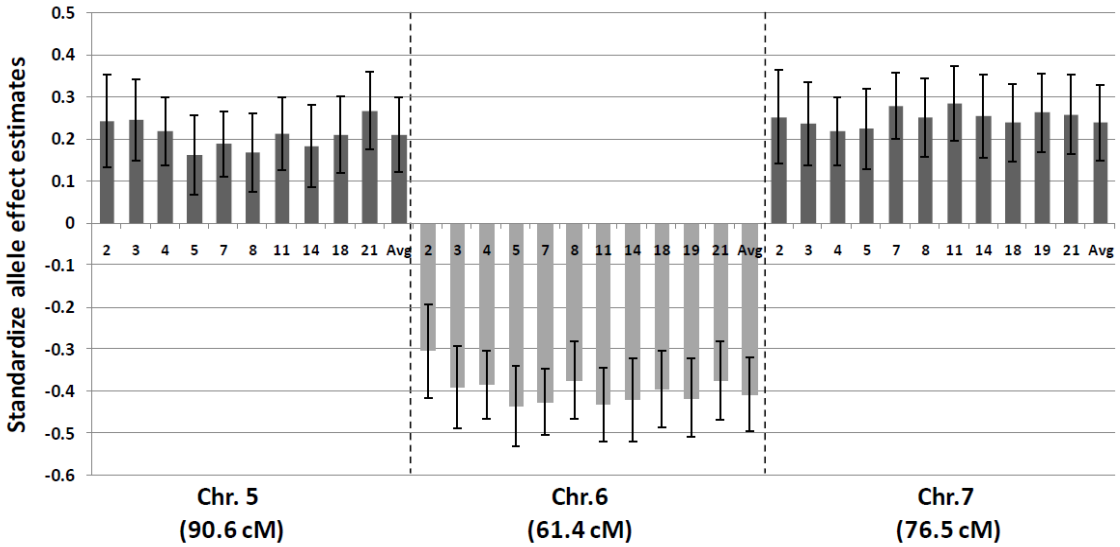


Figure 4.10. Standardized allele effect estimates are more consistent

Standardized estimated allele effects for scorers using the % scale at the same QTL as Fig.7a. Allele effects are shown as expected increase/decrease in disease from the MS71 allele. Units represent standard deviations on total variance for each individual scorer.

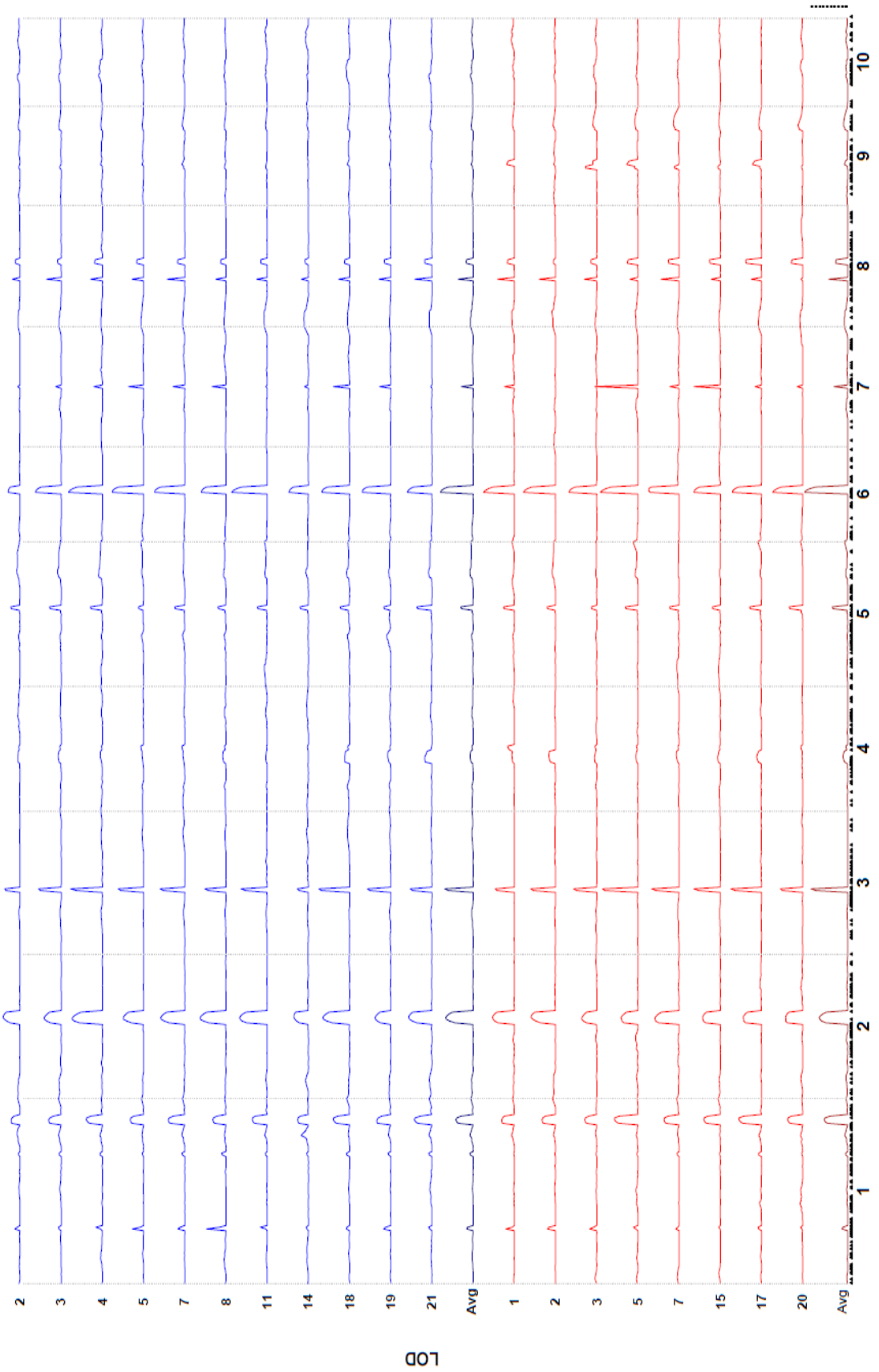
Figure 4.11. Inclusive composite interval mapping using different rating scales

Comparison of inclusive composite interval mapping (ICIM) for % and 0-9 scales. The BLUPs from the mixed model using data from all individuals for the two respective scales as well as a square root transformation of the % values were used for mapping. Shown on the x-axis is the genomic position in cM of the identified QTL. Chromosomes are marked by vertical gray lines and numbered accordingly. The positions of markers used in mapping are shown as black triangles. The QTL profile is shown as the log10 likelihood ratio, (LOD). The black and grey bar at right gives the scale of LOD = 20 with increments of 1.

Percentage
sqrt Percentage
0 to 9
1
2
3
4
5
6
7
8
9
10
Genomic Position
LOD

Figure 4.12 Inclusive composite interval mapping comparing different scorers.

Inclusive composite interval mapping was conducted for each individual scorer. Scorers are numbered to the left of each QTL profile. Ratings using the % and 0-9 scales are shown in blue and red, respectively. The x-axis is the genomic position in cM of the identified QTL. Chromosomes are marked by vertical gray lines and numbered accordingly. The positions of markers used in mapping are shown as black triangles. The QTL profile is shown as the log10 likelihood ratio, (LOD). The black and grey bar at right gives the scale of LOD = 20 with increments of 1.


LOD
2
3
4
5
7
8
11
14
18
19
21
Avg
1
2
3
5
7
15
17
20
Avg
1
2
3
4
5
6
7
8
9
10

To examine the effect of scorer variability and rating scale on genomic selection (GS) prediction, a ridge regression model (Meuwissen *et al.* 2001) was evaluated. The prediction accuracy was determined using a cross validation method by iteratively dropping one observation and then using the GS model from the remaining observations to predict the missing phenotype. Prediction accuracies were measured as the correlation between the predicted values and the empirical phenotypic observations. The prediction accuracies were in the range of 0.6 to 0.8. There was not a significant difference in the prediction accuracy of the % and 0-9 scales, despite the higher precision of the % scale.

DISCUSSION

Quantitative disease resistance remains an important area of research in plant breeding and phytopathology. To understand the underlying genetics of quantitative disease resistance, segregating populations have been used to map quantitative resistance loci (QTL). In virtually all of these studies mapping QTL, visual assessment has been used to determine disease severity. It remains unknown, however, how differences among scorers and rating scales affects the QTL mapping results. Based on the conclusions of these studies, various lines of follow-up are undertaken, including marker assisted selection of QTL in breeding programs and map-based cloning projects (both resource intensive endeavors). It is therefore pertinent to have a better understanding of how variability among individuals and the use of different rating scales affects QTL mapping results.

To examine the effect of individual variation and different rating scales

on QTL mapping, we utilized a population of 191 recombinant inbred lines from a cross between inbred lines MS71 and B73. Over 17,000 phenotypic observations from 22 different individuals on disease resistance were collected for this population, corresponding to over 80 disease ratings per line. This considerable data set allowed good inference regarding the consistency and the accuracy of visual evaluation of disease resistance in empirical field trials.

Disease rating experience leads to higher precision**.** We utilized the correlation between replications, the correlation between scorers, and the correlation to the mean of all observations as indicators of precision. The correlation between replications is based on the covariance between lines in the two respective replications. The covariance will decrease due to field/environmental effects and scorer error. Not accounting for field effects, lower scorer error will lead to a higher correlation between replications. There were considerable differences among scorers as well as between rating scales used. We used a simple 1-5 scale to measure scorer experience for general plant disease rating and also for NLB in maize. There was a significant increase in consistency for both NLB experience and the combined experience score and a slight trend for general experience. This shows that, as expected, experience scoring disease leads to higher consistency as measured by the correlation between replications. We tested the effect of experience on the consistency between different scorers. Again there was a significant effect of experience on the consistency, with additional experience by only one or both of the individuals increasing the correlation between the pair.

For a final examination of scorer experience, the average score of scorers with no previous NLB rating experience was compared that of the most experienced individual. It was found that the inexperienced average had

a higher correlation between replications than a single rating by a more experienced individual. This is analogous to the statistical phenomenon of increasing the sample size for a more precise estimate of the unknown sample mean. Though largely unfeasible, having multiple individuals score the same experiment will lead to a more precise rating. This is true even if the individuals are inexperienced.

Percentage ratings are more accurate than an ordinal scale**.** Comparison of scores using the 0-9 scale vs. the % scale showed that scoring with the direct percentage scale was significantly more precise than using the 0-9 scale. After accounting for scorer experience and rating time-point, the correlation between replications for scorers using the 0-9 scale was only 0.62 compared with 0.78 for scorers using the percentage scale. This held true after accounting for scorer experience. This is consistent with previous studies using computer simulated leaves that found higher accuracy when using a direct percentage scale than an ordinal scale (Hartung and Piepho 2007). The proposed reasons for this increased accuracy are several, and include difficulty in transferring from a percentage scale (observed) to an ordinal scale, estimation error compounded with rounding error imposed by the classes, and difficulty in maintaining consistency of which class is which. It was noted that individuals found the 0-9 scale easier to use but “vague” owing to small number of classes. We conclude that a direct percentage scale should be used when possible. If an ordinal scale is used, the scale should have at least twice as many increments as an individual is able to score. For example, if an individual can comfortably distinguish 10 different classes of disease levels, the scale should have at least 20 different levels to avoid compounding scoring error with rounding.

Precision can be increased through additional evaluations**.** The precision of the phenotypic values was increased through multiple ratings by different individuals. The increase in accuracy was likely due to a reduction of the scoring error as the average approached the true value for each line as the number of evaluations increased. This is the same statistical phenomenon of the law of large numbers in which the sample mean approaches the true mean with increasing sample size assuming unbiased sampling. The observed increase in accuracy is analogous to increasing the heritability of the trait. In breeding and agronomic applications, additional replications are typically added to achieve the same effect. While replications help to account for error due to field variability, there is limited utility of replications in reducing scoring error. Experience increased the correlation between replications presumably through reduction of scorer error. For a linear fit of the % ratings from replication 1 by replication 2 from the second rating of Year 2, the most experienced scorers had R^2 values of 0.76, 0.739, and 0.711 while less experienced scorers had an R^2 as low as 0.479. However the R^2 for the average of all individuals was 0. 804 and the average of inexperienced individuals was 0.783. This shows that, in this case, the heritability on a line mean basis of quantitative disease resistance can be increased with additional experience of the rater. An additional increase can be achieved by taking the average score of multiple different individuals. Depending on the crop and disease that is being evaluated, the plot size needed and the cost of field trials, an effective method for increasing the heritability of the evaluations would be to reduce the number of replications (e.g. from 3 or 4) and increase the number of individuals (e.g. from 1 to 4 or more) and number of evaluations that are conducted on each plot. This approach might be particularly effective

in crops with long generations or high costs (such as tree species).

Inclusive composite interval mapping produced highly consistent results**.** For QTL mapping, ICIM was highly robust to differences among individuals and rating scales in identifying the position and significance of QTL. As noted, the BLUPs using the data from all individuals should be the most accurate and precise assessment of actual disease levels. In this case, the ICIM mapping results from % and 0-9 ratings were almost identical (**Figure 4.11**). This indicates that for very precise phenotypic values the type of scale used should not have an effect on the position and significance of identified QTL. The consistency of the ICIM results from the % and 0-9 scales is in contrast to the non-linear relationship of these two ratings. The consistency of ICIM was further observed in the individual ratings. The same QTL were largely identified by all individuals, regardless of the rating scale used. The significance level of the identified QTL was affected by the accuracy of the ratings with less accurate ratings producing lower LOD values.

QTL allele effect estimates were highly variable**.** While the position of identified QTL was largely consistent, the estimated additive effect of those QTL was highly variable between scorers (**Figure 4.9**). At some QTL that were identified by all scorers, there was almost a threefold difference in the estimated effects. This variation in allele effect estimates is largely based on the population variance from the individual scorers. When allele effects were standardized, the estimated allele effects were more consistent though significant differences remained among individuals.

It is our observation that, for QTL mapping using visual observations of disease severity, the precision of the disease estimates has the greatest effect on power to detect QTL while the accuracy of estimations has the greatest

effect on the magnitude of estimated allele effects. Most previous studies mapping QTL should be considered fairly precise as multiple seasons and replications were generally used in the disease resistance evaluation. In this regard, the position of identified QTL can be considered reliable. However, the accuracy of these studies is not known (and cannot be known) and hence the estimated allele effects of identified QTL are likely specific to the individual study and the individual who conducted the rating.

REFERENCES

Balint-Kurti, P. J., J. C. Zwonitzer, M. E. Pe, G. Pea, M. Lee and A. J. Cardinal (2008). Identification of Quantitative Trait Loci for Resistance to Southern Leaf Blight and Days to Anthesis in Two Maize Recombinant Inbred Line Populations. *Phytopathology* **98** (3): 315-320.

Baumgarten, A., J. Suresh, G. May and R. Phillips (2007). Mapping QTLs Contributing to *Ustilago maydis* Resistance in Specific Plant Tissues of Maize. *Theoretical and Applied Genetics* **114** (7): 1229-1238.

Berman, M., M. L. Bason, F. Ellison, G. Peden and C. W. Wrigley (1996). Image Analysis of Whole Grains to Screen for Flour-Milling Yield in Wheat Breeding. *Cereal chemistry (USA)* **73** (3): 323-327.

Bock, C. H., P. E. Parker, A. Z. Cook and T. R. Gottwald (2008). Characteristics of the Perception of Different Severity Measures of Citrus Canker and the Relationships between the Various Symptom Types. *Plant Disease* **92** (6): 927-939.

Bock, C. H., P. E. Parker, A. Z. Cook and T. R. Gottwald (2008). Visual Rating and the Use of Image Analysis for Assessing Different Symptoms of Citrus Canker on Grapefruit Leaves. *Plant Disease* **92** (4): 530-541.

Bock, C. H., P. E. Parker, A. Z. Cook, T. Riley and T. R. Gottwald (2009). Comparison of Assessment of Citrus Canker Foliar Symptoms by Experienced and Inexperienced Raters. *Plant Disease* **93** (4): 412-424.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J. C. Glaubitz, M. M. Goodman, C. Harjes, K. Guill, D. E. Kroon, S. Larsson, N. K. Lepak, H.

Li, S. E. Mitchell, G. Pressoir, J. A. Peiffer, M. O. Rosas, T. R. Rocheford, M. C. Romay, S. Romero, S. Salvo, H. S. Villeda, H. Sofia da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich and M. D. McMullen (2009). The Genetic Architecture of Maize Flowering Time. *Science* **325** (5941): 714-718.

Campbell, K. G., C. J. Bergman, D. G. Gualberto, J. A. Anderson, M. J. Giroux, G. Hareland, R. G. Fulcher, M. E. Sorrells and P. L. Finney (1999). Quantitative Trait Loci Associated with Kernel Traits in a Soft x Hard Wheat Cross. *Crop Science* **39** (4): 1184-1195.

Hartung, K. and H. P. Piepho (2007). Are Ordinal Rating Scales Better Than Percent Ratings? A Statistical and “Psychological” View. *Euphytica* **155** (1): 15-26.

Joehanes, R. and J. C. Nelson (2008). Qgene 4.0, an Extensible Java QTL-Analysis Platform. *Bioinformatics* **24**: 2788-2789.

Li, H., G. Ye and J. Wang (2007). A Modified Algorithm for the Improvement of Composite Interval Mapping. *Genetics* **175** (1): 361-374.

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S. E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M. O. Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J. C. Glaubitz, M. Goodman, D. Ware, J. B. Holland and E. S. Buckler (2009). Genetic Properties of the Maize Nested Association Mapping Population. *Science* **325** (5941): 737-740.

Meuwissen, T. H. E., B. J. Hayes and M. E. Goddard (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **157** (4): 1819-1829.

Nita, M., M. A. Ellis and L. V. Madden (2003). Reliability and Accuracy of Visual Estimation of Phomopsis Leaf Blight of Strawberry. *Phytopathology* **93** (8): 995-1005.

Nutter, F. and P. Esker (2006). The Role of Psychophysics in Phytopathology: The Weber–Fechner Law Revisited. *European Journal of Plant Pathology* **114** (2): 199-213.

Nutter, F., P. Esker and R. Netto (2006). Disease Assessment Concepts and the Advancements Made in Improving the Accuracy and Precision of Plant Disease Data. *European Journal of Plant Pathology* **115** (1): 95-103.

Nutter, F. W. J., M. L. Gleason, J. H. Jenco and N. C. Christians (1993). Assessing the Accuracy, Intra-Rater Repeatability, and Inter-Rater Reliability of Disease Assessment Systems. *Phytopathology* **83** (8): 806-812.

O'Brien, R. D. and A. H. C. van Bruggen (1992). Accuracy, Precision, and Correlation to Yield Loss of Disease Severity Scales for Corky Root of Lettuce. *Phytopathology* **82**: 91-96.

Parker, S. R., M. W. Shaw and D. J. Royle (1995). The Reliability of Visual Estimates of Disease Severity on Cereal Leaves. *Plant Pathology* **44** (5): 856-864.

Pè, M. E., L. Gianfranceschi, G. Taramino, R. Tarchini, P. Angelini, M. Dani and G. Binelli (1993). Mapping Quantitative Trait Loci (QTLs) for Resistance to *Gibberella zeae* Infection in Maize. *Molecular and General Genetics MGG* **241** (1): 11-16.

Poland, J. A., P. J. Balint-Kurti, R. J. Wisser, R. C. Pratt and R. J. Nelson (2009). Shades of Gray: The World of Quantitative Disease Resistance.

Trends in Plant Science **14** (1): 21-29.

Wang, G. L., D. J. Mackill, J. M. Bonman, S. R. McCouch, M. C. Champoux and R. J. Nelson (1994). RFLP Mapping of Genes Conferring Complete and Partial Resistance to Blast in a Durably Resistant Rice Cultivar. *Genetics* **136** (4): 1421-1434.

Yu, J., J. B. Holland, M. D. McMullen and E. S. Buckler (2008). Genetic Design and Statistical Power of Nested Association Mapping in Maize. *Genetics* **178** (1): 539-551.

CHAPTER 5

Toward map-based cloning a gene that conditions quantitative resistance to northern leaf blight in maize

ABSTRACT

Despite the agronomic importance of quantitative disease resistance, little is known about its underlying genetic mechanisms. This is largely due to the small effects contributed by individual loci and the difficulty in obtaining accurate phenotypes. While advances have been made in association genetics for identifying genes and polymorphisms that are causally related to quantitative trait phenotypes, map-based cloning remains the most reliable and straight-forward method for identifying genes underlying quantitative trait loci (QTL) in plants. Recently, three genes that condition quantitative disease resistance have been identified using map-based approaches. Identification of these genes has shown diverse, previously unknown, types of genes involved in disease resistance. Here we advance toward identifying the underlying gene and causal polymorphisms for a quantitative resistance gene located on chromosome 1. Two concurrent approaches were used to converge on this locus of interest. An introgression of Tx303 in the background of B73 spanning 179.85 to 195.43 Mb had been previously associated with resistance and characterized in detail. Based on a large study with the maize nested association mapping population, a QTL was identified on Chr. 1 at 184.09 Mb on the B73 physical map. An F_2 population segregating for the Tx303 introgression was developed and 2,929 individual plants were genotyped to identify recombinant chromosomes in the introgression region. A sufficient number of recombinant lines were developed to allow positioning of the underlying causative polymorphism in a region of ~20-kb.

INTRODUCTION

In the search for genes underlying quantitative traits, map-based cloning in plants remains the most straightforward approach. Despite advances in genotyping, map-based cloning approaches remain time- and resource-intensive. However, the sequential dissection of a target introgression until the causal polymorphism can be localized to a single gene (or regulatory element) remains the 'gold-standard' for gene identification in plants. Association mapping is one alternative approach to gene identification for a quantitative trait of interest. Using diverse inbred collections, association mapping can be applied to identify genes that condition the quantitative phenotype. With the low linkage disequilibrium found in maize, association mapping can provide gene level resolution (Flint-Garcia *et al.* 2005). However, one caveat to association mapping is that, in the absence of high density SNP markers (>>1M in maize), information on candidate genes must be available to target for association mapping. With very little known about quantitative disease resistance, this is not currently possible. Targeted gene silencing can also be used to associate a gene with a quantitative phenotype. This approach has been successfully applied in rice to link quantitative resistance with a family of germin-like genes (Manosalva *et al.* 2009). As with association mapping, however, previous knowledge of candidate genes is needed. A completed genome sequence is also advantageous, which until recently was not available in maize.

Recently, map-based cloning approaches have been used to identify three genes controlling quantitative disease resistance (QDR). In wheat, rust resistance gene *Lr34* was mapped to a 0.15-cM target region of 363-kb using three different populations with a total of 4,032 individuals. *Lr34* was identified

as a putative ABC-transporter (Krattinger *et al.* 2009). The stripe rust resistance gene *Yr36* has been isolated by map-based cloning and found to encode a Kinase-START protein. Using a population of 4,500 F_2 plants, *Yr36* was narrowed to a 0.14-cM region of 314-kb. Further addition of markers narrowed the region to 0.02-cM which contained two candidate genes (Fu *et al.* 2009). The rice gene *pi21* was narrowed to a region of 1.7-kb using two populations with a total of 3,717 individuals. *Pi21* was identified as a proline-rich gene of unknown function (Fukuoka *et al.* 2009). A recessive mutation in this gene causes the *pi21* resistance phenotype. The fourth QTL that has been isolated to date is *Rcg1* from maize. *Rcg1* confers resistance to *Colletotrichum graminicola* and encodes a nucleotide binding leucine-rich repeat (NB-LRR) type gene (Broglie *et al.* 2006; Wolters *et al.* 2006). These reports show that map-based cloning can be successfully used to identify the genes responsible for quantitative disease resistance. With the exception of *Rcg1*, the QDR genes that have been cloned to date represent new types of genes not previously known to be associated with disease resistance in plants.

Through screening a BC_3S_3 Tx303 introgression library in the B73 background, Chung *et al.* (submitted 2009) previously identified a line with an introgression in maize bin 1.06 that conferred increased resistance to northern leaf blight (NLB). Analysis of segregating material from the backcross line confirmed that resistance was associated with the Tx303 allele at 1.06 and that the resistance allele was partially dominant. Detailed characterization of the Tx303 resistance allele indicated that the resistance was most effective early in pathogenesis by reducing infection efficiency (C. Chung *et al.*, submitted). From concurrent work on the maize nested association mapping (NAM) population, a QTL was identified at the same location as the

introgression identified by Chung *et al.* Based on the confirmed effect of the Tx303 introgression in bin 1.06, the well-characterized and unique mode of resistance, and the suitable genetic resources developed through previous studies (C. Chung *et al.* submitted), this QTL is a good target for further study and map-based cloning. The information from NAM can lend additional inference in positioning, characterizing, and confirming genes indentified in the map-based cloning approach. Identification of a gene conferring quantitative resistance to NLB in maize will not only assist furthering the understanding of molecular mechanisms of QDR but also in furthering resistance breeding efforts for this economically important disease.

MATERIALS AND METHODS

Plant Materials: A near-isogenic line carrying a Tx303 introgression in maize bin 1.06 was previously identified as more resistant than the recurrent parent B73 (C.Chung *et al.* in prep). An F_2 population (38_19E) was developed by crossing the isogenic line to B73. Segregation analysis confirmed that the introgression on Chr. 1.06 conferred resistance to NLB. From this population, 15 heterozygous individuals were selected to form a larger mapping population. A population of 4,080 F_2 seeds was planted in 2009. Individual plants from this population were genotyped with three SNP markers; two flanking markers for the introgression and a marker in the middle of the introgression (**Table 1, dark blue**).

Table 5.1 Markers used for characterization and fine-mapping.

recombinants															
kb/rec															
cM/Mb															
cM in NILs															
BIN	1.01	1.03	1.03	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.06
NAM marker															
physical_AGI				162.44	165.26	168.28	168.43	168.51	168.94	169.31	169.31	169.31	170.69	?	175.50
NIL familiy	umc1071	umc2383-M13	bnlg439	umc1906	umc1603	snp_01_0016	snp_01_0017	chr1|TX303|168512062|A/(	chr1|TX303|168938311|A/(	chr1|TX303|169305992|T/(	chr1|TX303|169305984|C/A	chr1|TX303|169305984|C/T	umc1386	bnlg1270	an1_5
38_19E	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B
38_17A	B	B	B	B	B								B	B	
38_15G	B	B	B	B	B								B	B	
				ssr_01_067	ssr_01_082	snp_01_0016	snp_01_0017	snp_01_0022	snp_01_0037	snp_01_0038	snp_01_0040	snp_01_0039	ssr_01_078	ssr_01_079	snp_01_0012
BAC															
NOTES															

recombinants									21			74		36	
kb/rec									58			14		18	
cM/Mb						-			0.29			0.90		1.54	
cM in NILs						0			0.36			1.26		2.24	
BIN	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06
NAM marker											m84				
physical_AGI	176.90	178.15	176.89	178.50	179.85	180.26		180.57	181.47	181.80	182.39	182.47	182.84	183.11	183.54
NIL familiy	chr1_AC193489_14841	chr1_AC216194_7076	chr1_AC193489_13202	ch1_AC191428_15886	PZA00455.16	ch1_AC202158_78820	ch1_AC202158_78773	umc1754	chr1|181473188|A/G	chr1|181804065|G/A	PZA02191.1	chr1|182474151|A/G	chr1|182841737|C/T	chr1|183107525|C/T	chr1|183536439|A/C
38_19E	B	B	B	B	H?	H	H	H	H	H		H	H	H	H
38_17A						H		H							
38_15G						H		H							
	snp_01_0046	snp_01_0045	snp_01_0044	snp_01_0043	snp_01_0011	snp_01_0042 C2	snp_01_0041	ssr_01_027	snp_01_0054 C2	snp_01_0053 C2	snp_01_0001	snp_01_0056 C1	snp_01_0052 C1	snp_01_0057 C1	snp_01_0058 C2
BAC	AC193489	AC216194	AC193489	AC191428	AC197435.2	AC202158	AC202158		AC196957.3	AC213369.4	AC186606.4	AC186606.4	AC210429.3	AC186228.4	AC194176.3
NOTES					hemizygous						monomorphic				

recombinants		64		43	3		4						435			
kb/rec		11		16	10		259						23			
cM/Mb		1.51		1.07	1.78		0.07						0.75			
cM in NILs		3.33		4.06	4.11		4.18						11.61			
BIN	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	1.06	3.02	5.00
NAM marker	m85	m86	m87													
physical_AGI	183.65	183.83	184.09	184.52	184.55		185.58			193.64			195.43			0.66
NIL familiy	PHM1968.22	PZA00068.1	PHM5622.21	chr1|184516928|G/A	chr1|184545721|A/G	umc1811	chr1|185582569|A/C	umc2234	umc1919	chr1|193635759|A/G	umc1668	umc1035	PZA00619_3	umc1335	bnlg1647	mmc0151
38_19E		H	H	H	H	H	H	H	H	H	H	H	H	B		T
38_17A						H	H	H	H	H	H	B	B	B	B	T
38_15G						H	H	H	H	H	H	H	H	B	B	T
		snp_01_0015 C2	snp_01_0014	snp_01_0059 C2	snp_01_0060 C2	ssr_01_069	snp_01_0047 C2	ssr_01_028	ssr_01_071	snp_01_0062 C2	ssr_01_072	ssr_01_073	snp_01_0005			
BAC	AC185412.3	AC177818.4	AC183930.4	AC187262.4	AC187262.4					AC216732.3						
NOTES	monomorphic															

The maize nested association mapping population (NAM) was previously analyzed for NLB resistance (Chap. 2). Tx303 is one of the 25 diverse inbred lines that were used as parents for the NAM families. In an effort to gain further insight into the QTL on Chr. 1.06, the Tx303 family was analyzed separately here. The Tx303 family consisted of 188 individuals that were evaluated for NLB resistance over three years.

DNA Extraction and Genotyping: For genotyping the fine-mapping population, a piece of seedling leaf tissue approx 1 mm^2 was collected and placed in a 0.2-ml 8-strip PCR tube. Extract-N-Amp^{TM} kit (Sigma-Aldrich Co.) was used for a 'quick and dirty' DNA extraction. Samples were briefly centrifuged at 4000 rpm to get the leaf samples to the bottom of the tube. 8-µl of extraction solution was added and samples were held at 95^{o}C for 10 min. 8-µl of the dilution solution was then added and samples were stored at 4^{o}C. Samples were transferred to 96 well plates and diluted 1:100 in water. Using the Biomek robot (Cornell Core Facility), 10-µl was then arrayed into 384 PCR plates (Kbioscience, Hoddesdon Herts, UK), centrifuged, and dried. Dried plates were then held at 4^{o}C until use in PCR.

SNP genotyping was conducted using KASPar (Kbioscience, Hoddesdon Herts, UK). To successfully genotype on the 'quick and dirty' DNA samples, a 1:100 dilution was used as specified above to reduce the amount of contaminants in the PCR reaction. A 4-µl reaction was prepared according to the KASPar manual and added on-top of the dried DNA. The reaction mix consisted of 2-µl of KASPar reaction mix, 0.055-µl of primer assay mix, 0.032-µl $MgCl_2$ (50mM), and 2-µl of H_20. Due to the low DNA concentration in the reaction additional cycles were added to the second step of the thermo-cycling

program. The PCR program was as follows: 94^{o}C for 15 minutes; 20 cycles of 94^{o}C for 10 seconds, 57^{o}C for 5 seconds, and 72^{o}C for 10 seconds, followed by 26 cycles of 94^{o}C for 10 seconds, 57^{o}C for 20 seconds, and 72^{o}C for 40 seconds.

Based on the genotyping results, recombinant plants were identified from the population. To confirm the recombination events, a new sample of leaf tissue was collected and lyophilized. DNA was extracted from this tissue in 2-ml 96 well plates using the CTAB method. Genomic DNA was diluted 1:50 and arrayed into 384 well plates and dried as described above. KASPar SNP genotyping was conducted as described using 18 cycles for the second round of cycling in the program.

SNP marker development: SNP markers were developed using SNPs from the first generation maize haplotype map (Gore *et al.* 2009) (http://www.maizegenetics.net/maize-hap-map). SNPs polymorphic between B73 and Tx303 were identified and sequence surrounding the SNP was obtained from the B73 reference sequence (http://www.maizesequence.org). SNP primers were developed using the on-line KASPar primer design utility from Kbioscience (http://213.123.130.96/primer-picker.htm). Primers were tested on B73 (x4), Tx303 (x4), and a sample of 16 F_2 individuals from the segregating population to confirm polymorphism in the NIL.

Linkage mapping: Inclusive composite interval mapping (ICIM) was conducted for the Tx303 x B73 RIL population using qGene v4.3.2 (Joehanes and Nelson 2008). For ICIM, forward cofactor selection was used with the default threshold. Interval mapping was done with a 1-cM step interval. To compare the position of a co-localizing QTL in NAM, individual markers across the region were evaluated. A joint linear model was fit for all QTL previously

identified in NAM (Chap. 2). A marginal F-test was used to assay the significance of the QTL marker in bin 1.06. This marker was then dropped from the model and flanking markers were iteratively assayed for significance based on the same test. The significance of each marker was assayed using the marginal F-test for improvement of the linear regression model.

RESULTS

From large-scale analysis of resistance to northern leaf blight (NLB) in maize using the nested association mapping population, a quantitative trait locus (QTL) for NLB resistance was identified on Chr. 1 at 184.09 Mb on the B73 AGI physical map (see Chapter 2). The 95% confidence interval for this QTL fell between 97.7 and 100.3 on the NAM genetic map, which corresponds to 182.39 Mb and 185.33 Mb on the B73 physical map (AGP 4a.53) (**Figure 5.1**). Estimated allele effects for each of the founder lines indicated that there were two classes of alleles and that resistance was segregating in nine of the 26 NAM families. Considering the genetic design of the NAM population, this type of locus should have high power for nested association mapping.

In an attempt to capture more information from the available data on NAM, the Tx303 family was analyzed separately. Inclusive composite interval mapping (ICIM) was conducted to identify QTL in the Tx303 x B73 family (**Figure 5.2**). Five QTL were identified, including a QTL on Chr. 1.06. Consistent with the findings of C. Chung *et al.* (submitted, 2009), the Tx303 allele conditions resistance at this position. The QTL position identified by NAM was approximately 5 Mb away from, and outside of the confidence interval of the QTL position identified by ICIM in the Tx303 family.

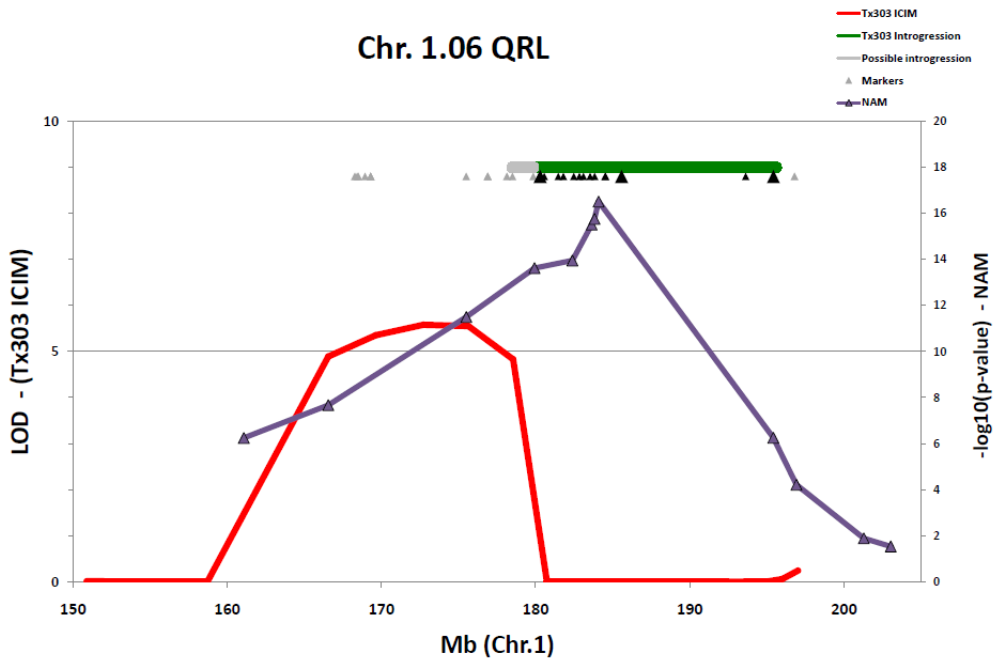


Figure 5.1. Linkage mapping in NAM and the Tx303 x B73 family compared to the Tx303 NIL introgression

Region of maize chromosome 1 shown in Mb (B73 AGP 4a.53). The region of Tx303 introgression associated with resistance to NLB is shown in green. A region of potential introgression where the end of the introgression was mapped is shown in gray. The LOD profile for inclusive composite interval mapping in the Tx303 x B73 population is shown in red. Single marker regression for NAM (accounting for background QTL) is shown in purple with marker positions dented with triangles. The positions of markers used for fine-mapping are shown in black. The markers shown in gray were not polymorphic in the Tx303/B73 NIL.

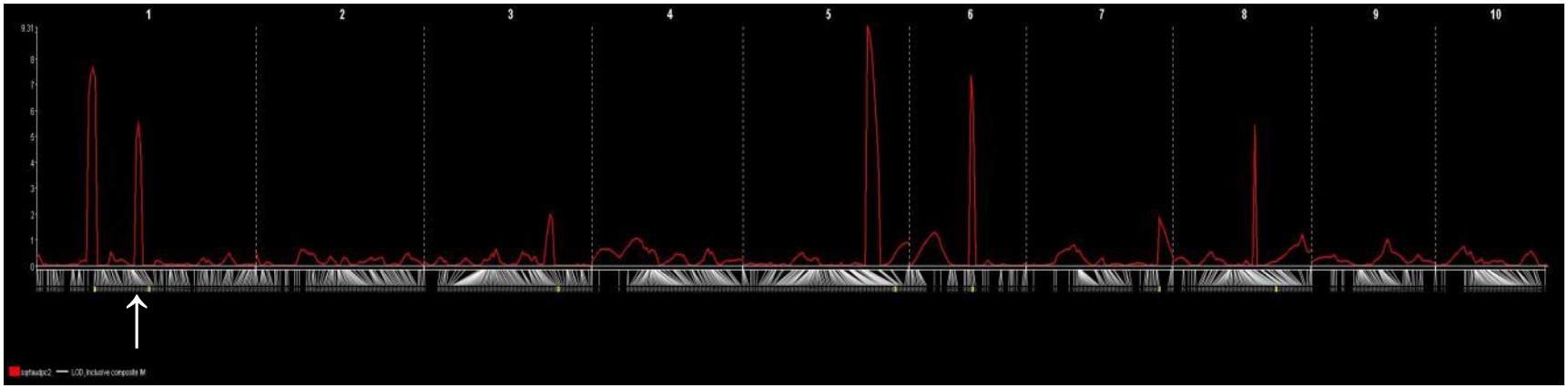


Figure 5.2. Inclusive composite interval mapping in Tx303 x B73 RIL population

LOD profile for inclusive composite interval mapping in Tx303 x B73 RIL population. The second QTL peak on Chr.1 (shown with arrow) corresponds to bin 1.06 at the location of previously identified QTL in Tx303.

The QTL identified by NAM could be a composite of two or more closely linked QTL in the individual families as NAM represents multiple different families (and alleles). Alternatively, the position identified by interval mapping in the Tx303 family could be incorrect due to limited population size and/or an additional QTL in the region.

There are 1106 marker genotypes on NAM for an average marker distance of 1.2-cM. While this is a very high marker density for a typical mapping population, further resolution could be obtained from NAM with additional markers. NAM lines with recombination events in the region were identified for marker saturation. Eight of the SNPs designed for fine-mapping in the Tx303/B73 NILs were evaluated on 600 NAM lines. Despite robust genotyping results using Tx303 and B73 derived material, all but one of the eight markers failed for NAM. Secondary polymorphisms at the SNP primer sites are likely the cause of failure. While marker saturation of NAM is still a valid approach for fine-mapping QTL, it was not successful in this study due to failure of the SNP markers.

To better characterize the position of the Tx303 introgression in the NIL (38_19E), SNP markers were added to the introgression region. Based on markers that were polymorphic between Tx303 and B73 and segregating in the NIL-F_2 population, the ends of the Tx303 introgression were mapped between 178.6 – 180.2 Mb for the left and between 195.43-196.92 Mb on the right (**Table 5.1, Figure 5.1**). The QTL region identified in NAM directly corresponds to the Tx303 introgression in the NIL. However, the QTL region identified by ICIM in the Tx303 family was outside of the introgression (save a slight potential overlap at the left end of the introgression). It was not clear why the QTL position identified by ICIM in the Tx303 x B73 family did not

correspond to the same position as the NIL introgression. Assuming that there was not a small secondary introgression in the region, it appears that the QTL position identified by ICIM was incorrect.

In the NIL-F_2 population, 3,166 plants were genotyped with three SNP markers (Table 1). Reliable genotype results were obtained for 2,929 of these plants. From this population, 273 and 601 plants were found to have recombinant genotypes in the two intervals between the left flanking marker (ch1_AC202158_78820), a central marker (chr1|185582569|A/C) and the right marker (PZA00619.3), respectively. Based on the results from NAM, the left interval was targeted for further study. To confirm the recombination events in this interval, tissue was re-collected from the plants that had recombinant genotypes. Of the 267 plants examined, 234 were confirmed as recombinants. For this validation, about 12% of the original genotypes were found to be incorrect. This percentage is likely higher than the actual error rate, as recombinants and possible recombinants (i.e. plants with questionable genotyping results) were tagged for confirmation. From the original population, ~93% of the plants had reliable genotyping results.

From the original F_2 population, 100 plants were identified with recombinant genotypes. During the summer of 2009, progeny from these plants were grown to select homozygous individuals for selfing. In addition to selecting homozygous plants, further recombinant genotypes were also identified. For the 100 recombinant lines, 1,656 plants were genotyped, averaging 16 per line. Two lines were identified as non-recombinant (mis-genotyped in the previous generation). Homozygous individuals from the remaining lines were identified and self-pollinated. In the first marker interval, 54 recombinant plants were identified and 40 recombinant genotypes were

identified in the second interval. Tissue was re-collected from individuals with recombinant genotypes for the first interval. These plants were re-genotyped to confirm the recombinant genotypes.

From individuals identified with recombinant genotypes in the first interval from the fine mapping population and from the recombinant $F_{2:3}$ lines, 240 individuals were selected to be advanced. These lines had been confirmed as recombinant genotypes and had sufficient seed for advancement. The seed was sent to winter nursery in Argentina where ~16 plants per line will be self pollinated and shelled individually. Seed from each individual will be germinated in the greenhouse and bulked for DNA extraction. Lines will be genotyped to identify recombinant homozygous lines. Recombinant lines will be planted in Aurora 2010 for phenotypic evaluation for NLB resistance.

DISCUSSION

Map-based cloning approaches have been successfully used in plants to identify genes underlying QTL. The same approach has also been used to identify several genes for QDR in plants. Although this approach is demanding with regard to resources, time, consumables, and labor, map-based cloning remains a direct and reliable method for identifying genes underlying QTL. Here we evaluated a segregating population of 2,929 individuals as well as 1,656 individuals from 100 lines for fixing recombinant genotypes for a QTL which conditions resistance to NLB. From the fine-mapping population, 874 plants were identified with recombinant genotypes. From the 100 lines previously identified as recombinant genotypes, homozygous recombinant individuals from 98 of the lines were identified and self-pollinated. An addition

94 recombinant plants were also identified. From the fine-mapping population, 240 individuals were re-genotyped, confirmed as recombinants, and advanced to winter nursery as lines for fixing.

Based on concurrent work in NAM, focus was placed on the left portion of the introgression. In this interval, 240 recombinant genotypes are being advanced to homozygous lines in winter nursery. An additional 25 recombinant lines were fixed in the summer of 2009. Together these lines will provide a population of 265 recombinant genotypes. The physical size of this interval is 4.18 Mb (B73 AGP 4a.53). With 265 recombination events, it is expected that there will be a cross-over every 15.5-kb on average. This resolution should be sufficient to narrow the QTL to a single gene in maize. It is expected that additional recombination events will be identified during genotyping of the lines that are being fixed. An average of 16 individual plants will be genotyped for these lines. As these lines are not segregating for the full introgression, half of the 3,840 plants genotyped are expected to have informative recombinant genotypes, further increasing the resolution for this map-based cloning of the NLB resistance QTL.

In an attempt to triangulate available information, linkage-based positioning of QTL from NAM and the Tx303 x B73 family was used to narrow the QTL region for NLB resistance at 1.06. The QTL position identified by NAM was consistent with the Tx303 introgression and narrowed the QTL region to half of the introgression. Mapping QTL using ICIM in the Tx303 x B73 family resulted in identification of five QTL. Consistent with previous observation in the NILs, a QTL was mapped in bin 1.06 with resistance conferred by the Tx303 allele. However, with marker saturation of the NIL introgression, the region identified by ICIM did not overlap with the region of

introgression. It is not apparent why the QTL position identified by linkage mapping in the Tx303 x B73 RIL population was not consistent with the known position of the QTL in the NILs. A secondary QTL in the region could be influencing the position of the target QTL. However, if inference was based only on the ICIM results it appears that the incorrect region would be identified for follow-up studies.

REFERENCES

Broglie, K. E., K. H. Butler, C. A. De Silva, T. J. Frey, J. A. Hawk, D. S. Multani, C. Wolters and J. C. Petra (2006). Polynucleotides and Methods for Making Plants Resistant to Fungal Pathogens. United States, E.I. du Pont de Nemours and Company, Pioneer Hi-Bred International, Inc.; University of Delaware. **United States Patent 20060223102**.

Flint-Garcia, S. A., A.-C. Thuillet, J. Yu, G. Pressoir, S. M. Romero, S. E. Mitchell, J. Doebley, S. Kresovich, M. M. Goodman and E. S. Buckler (2005). Maize Association Population: A High-Resolution Platform for Quantitative Trait Locus Dissection. *Plant Journal* **44**(6): 1054-1064.

Fu, D., C. Uauy, A. Distelfeld, A. Blechl, L. Epstein, X. Chen, H. Sela, T. Fahima and J. Dubcovsky (2009). A Kinase-Start Gene Confers Temperature-Dependent Resistance to Wheat Stripe Rust. *Science* **323**(5919): 1357-1360.

Fukuoka, S., N. Saka, H. Koga, K. Ono, T. Shimizu, K. Ebana, N. Hayashi, A. Takahashi, H. Hirochika, K. Okuno and M. Yano (2009). Loss of Function of a Proline-Containing Protein Confers Durable Disease Resistance in Rice. *Science* **325**(5943): 998-1001.

Gore, M. A., J.-M. Chia, R. J. Elshire, Q. Sun, E. S. Ersoz, B. L. Hurwitz, J. A. Peiffer, M. D. McMullen, G. S. Grills, J. Ross-Ibarra, D. H. Ware and E. S. Buckler (2009). A First-Generation Haplotype Map of Maize. *Science* **326**(5956): 1115-1117.

Joehanes, R. and J. C. Nelson (2008). Qgene 4.0, an Extensible Java QTL-Analysis Platform. *Bioinformatics* **24**: 2788-2789.

Krattinger, S. G., E. S. Lagudah, W. Spielmeyer, R. P. Singh, J. Huerta-Espino, H. McFadden, E. Bossolini, L. L. Selter and B. Keller (2009). A Putative ABC Transporter Confers Durable Resistance to Multiple Fungal Pathogens in Wheat. *Science* **323**(5919): 1360-1363.

Manosalva, P. M., R. M. Davidson, B. Liu, X. Zhu, S. H. Hulbert, H. Leung and J. E. Leach (2009). A Germin-Like Protein Gene Family Functions as a Complex Quantitative Trait Locus Conferring Broad-Spectrum Disease Resistance in Rice. *PLANT PHYSIOLOGY* **149**(1): 286-296.

Wolters, P., T. Frey, A. Conceição, D. Multani, K. Broglie, S. Davis, K. Fengler, E. Johnson, K. Bacot, K. Simcox, T. Weldekidan and J. Hawk (2006). Map Based Cloning of a QTL for Anthracnose Stalk Rot Resistance in Maize. paper W412, Plant and Animal Genome Meeting San Diego.

CHAPTER 6

Phenotypic evaluation of alleles under selection in maize

ABSTRACT

Selection mapping is a method for identification of QTL for important agronomic traits by examining changes in allele frequency over cycles of selection. We applied selection mapping in a diverse maize population that was improved through recurrent selection for resistance to northern leaf blight (NLB) to identify loci under selection that putatively associate with northern leaf blight resistance. The population *per se* was also evaluated for resistance to southern leaf blight (SLB) and gray leaf spot (GLS). Although not selected for SLB or GLS resistance, the population had increased resistance for these diseases over the cycles of selection. The caveat of this selection mapping, however, is that identification of significant markers could be associated with any number of selected traits as the populations examined were improved for secondary traits of agronomic interest. In this study we developed two recombinant inbred line (RIL) populations with the objective of confirming the phenotypic effect of selected loci. The RIL populations were phenotyped for NLB, SLB and GLS resistance during the summer of 2009. Positive correlations between resistances were observed, indicating co-localizing QRL and/or pleiotropic MDR genes. Phenotypic correlations were also examined for relative maturity and disease resistance. Consistent with previous results from these pathosystems, a negative correlation between maturity and disease resistance was observed for SLB and GLS. However, NLB resistance in both populations and SLB resistance in one of the RIL populations were positively correlated with relatively maturity. This finding is unique in the populations studied to date for resistance to NLB and SLB.

INTRODUCTION

Selection mapping is a method for identification of quantitative trait loci (QTL) through examining allele frequency changes in populations that have undergone selection. This approach can be applied to populations under both natural and artificial selection (Coque and Gallais 2006; Wisser *et al.* 2008). By comparing allele frequency changes during the generations of selection to that expected from drift, loci under positive selection can be identified. Selection mapping has several advantages over alternative QTL mapping strategies. Selection mapping can be conducted directly on material from a breeding population, eliminating the time and resources needed for development of dedicated genetic stocks for study of quantitative traits. The caveat of selection mapping, however, it that putatively selected alleles are identified independently of phenotypic observations. In order to validate the selection mapping approach, secondary genetic experiments must be conducted to determine trait-marker associations at the putatively selected loci.

In our group, selection mapping was previously employed to successfully identify markers associated with quantitative trait loci (QTL) for northern leaf blight (NLB) resistance in maize (Wisser *et al.* 2008). A diverse maize population that had undergone four cycles of full-sib S_1 recurrent selection was assayed with genome-wide SSR markers to identify loci that had exceeded the expected frequency changes that could be accounted for by drift. Using a simulation of the recurrent selection program to determine the expected changes possible through drift, 25 loci were found to have allele changes beyond the expectation. One locus on Chr. 8 where several putatively selected loci were located was evaluated in an F_2 derived from the

RS population. Segregation analysis showed that the favorable allele from the RS population was associated with resistance. To further demonstrate the utility of selection mapping, we have developed two larger recombinant inbred populations from the same RS material. The larger advanced populations will permit greater power for detection of loci of small effect along with additional replication and evaluation for resistance to other maize diseases of interest.

MATERIALS AND METHODS

Recurrent selection population: The recurrent selection (RS) population, Pool 30 (P30), used in this study is described in detail previously (Ceballos *et al.* 1991; Wisser *et al.* 2008). P30 is one of eight diverse maize populations from CIMMYT that was improved for northern leaf blight resistance through four cycles of full-sib, S_1 recurrent selection. P30 is described as an early sub-tropical yellow dent. The seed for P30 provided by CIMMYT was derived from the selected population by two separate seed increases of approximately 100 x 100 and 60 x 60 random plant-to-plant pollinations. The seed from these increases was bulked in equal quantities. The germplasm is documented at CIMMYT as: Pool 30 C0 (N)/TL2001B-6153-176; Pool 30 C1 E.T./TL2001B-6152-29; Pool 30 C2 E.T./TL2001B-6152-30; Pool 30 C3 E.T./TL2001B-6152-31; Pool 30 C4 E.T./TL2001B-6152-32. From the CIMMYT germplasm stock, an increase of Pool 30 cycle 0 (#176) and Pool 30 cycle 4 (#32) were derived by sib-mating ~46 individuals from each of the respective cycles and bulking equal quantities of seed. The seed from this increase was only used for phenotypic evaluations and not any genetic studies. A second recurrent selection population, Pool 29 (P29), was developed concurrently with P30 (Ceballos *et al.* 1991). P29 is also an early

sub-tropical population, but composed of yellow flint material. Seed for P29 was increased as described for P30, and used for evaluation of the population *per se*.

Recombinant Inbred Line Populations: Two recombinant inbred line populations (hereafter referred to as Pop5 and Pop10) were derived from a single individual (#32-24) from the last cycle of selection in P30 crossed with the inbred line B73. Two individual F_1 plants were selfed (#32-24-5 and #32-24-10) to make F_2 populations. For Pop5, 330 F_2 seed were planted in Aurora, NY during the summer of 2007 and all individuals were selfed. For Pop10, 400 F_2 seed were planted in Homested, FL in 2007-2008 and all individuals were selfed. $F_{2:3}$ lines were planted in Aurora 2008 as single rows. For each line, two plants were selfed. $F_{3:4}$ lines with sufficient seed were planted in Homestead, FL in 2008-2009 and two plants were selfed. The resulting $F_{4:5}$ lines were evaluated for NLB, SLB and GLS resistance during the summer of 2009. Seed for the RILs was increased in Aurora, NY 2009 by sib-mating 8 to 12 plants per line. The final numbers of $F_{4:5}$ RILs for Pop5 and Pop10 were 197 and 241, respectively.

Back-cross materials: From the original F_1 generation of 32-24 x B73, 10 individuals were selected to advance by back-crossing to B73. From the nine F_1 individuals, 25-30 independent BC_1 lines were developed. One of the F_1 individuals used for backcrossing was 32-24-10, the same individual used for RIL population development. Sixty independent BC_1 lines were developed for this pedigree. All BC_1 lines were crossed to B73 and advanced to the BC_2 generation with 2-3 BC_2 lines for each BC_1. For future genetic studies, >300 BC_2 lines were derived from the original RS plant used for RIL population development.

Phenotypic evaluation

Northern Leaf Blight: Evaluation of RILs for northern leaf blight (NLB) resistance was conducted at the Cornell Musgrave Research Farm during the summer of 2009. Lines were planted in augmented incomplete blocks with two replications. Each block consisted of 16 RILs and B73 inbred line as a common check. Plants were inoculated with *E. turcicum* race 1 (Isolate NY001 – R. Nelson Lab) at the 6 to 8 leaf stage. Disease severity ratings were conducted at three times at 10-day intervals. Disease severity was evaluated as percentage total leaf area in a row covered with necrotic NLB lesions. A mixed model repeated measures analysis was run in SAS software (Proc Mixed) to get best linear unbiased predictions for each line (BLUPs).

To confirm the resistance in P29 and P30 populations *per se*, samples of cycle 0 and cycle 4 from each population were also evaluated in 2009. Six replications of P29-C0, P29-C4, P30-C0 and P30-C4 along with B73 and Mo17 for inbred checks were planted, inoculated and rated at two time points for disease. A simple linear model was fit and LS mean effects for each of the cycles were compared.

Southern Leaf Blight: A sample of the population *per se* of Pool 29 and Pool 30 Cycles 0 and 4 were evaluated in Clayton, NC during the summers of 2007 and 2008 for resistance to southern leaf blight (SLB). Trials were planted on 4/24/2007 and 4/22/2008, respectively. The trials were arranged in a randomized complete block design with three blocks in 2007 and three blocks in 2008. Each block consisted of one row of Pool 29 C0, Pool 29 C4, Pool 30 C0, Pool 30 C4, B73 inbred, and Mo17 inbred, with ~10 plants in each row. The trials were inoculated with *Cochliobolus heterosrophus* race 0 (isolate 2-16Bm from P. Balint-Kurti lab, NC State Univ.) on 5/31/2007 and

6/9/2008, respectively. The trials were rated on a row basis for resistance to SLB using a 1 to 9 scale with 9 representing no disease and 1 being completely blighted. The trial was evaluated at three time points spaced 12 days apart in 2007 and six time points spaced seven days apart in 2008. Area under the disease progress curve was calculated for 2007 and 2008 data. To determine the consistency of phenotypic observations across environments, the two years were analyzed separately. Block was considered a random effect and pedigree (e.g. Pool29 C0) was considered a fixed effect. LS means were calculated and used to compare the different cycles of selection.

The RIL population was evaluated for SLB resistance in 2009 using the same experimental conditions described for the populations *per se*. Lines were planted in augmented incomplete blocks with two replications. Each block consisted of 16 RILs and B73 inbred line as a common check. Ratings were conducted at three time-points. A mixed model repeated measures analysis was run in SAS software (Proc Mixed) to get best linear unbiased predictions for each line (BLUPs).

Gray leaf spot: Gray leaf spot (GLS) was evaluated on the population *per se* in Andrews, NC, during the summers of 2007 and 2008. Trials were arranged as described for SLB. In 2007, GLS ratings were conducted at two time-points using a 1 to 9 scale where 9 is absence of disease and 1 is completely blighted. In 2008, the GLS trials were evaluated at four time-points. Area under the disease progress curve was calculated separately for 2007 and 2008 data. Block was considered a random effect and pedigree was considered a fixed effect. LS means were calculated and used to compare the different cycles of selection.

GLS resistance was evaluated under natural disease pressure on the

RIL population in Blacksburg, VA during the summer of 2009. Trials were planted in augmented incomplete blocks with two replications as described above. Ratings were conducted at three time-points using a 1 to 5 rating scale where 1 is the absence of disease and 5 is completely blighted. A mixed model for repeated measures was used to account for block and replication effects.

RESULTS

Evaluation of populations *per se*: Evaluation of a sample from cycle 0 and cycle 4 from P29 and P30 in NY during the summer of 2009 confirmed a large increase in resistance to NLB through recurrent selection (**Figure 6.1**). To determine any potential multiple disease resistance, the populations *per se* were evaluated for SLB and GLS resistance. There was significant improvement for resistance to SLB in P30 between cycle 0 and cycle 4. The increased resistance in the last cycle of selection was consistent across the two environments evaluated for SLB resistance in the population *per se* (**Figures 6.2a and b**). There was a slight but non-significant difference in resistance in P29 (p-value = 0.087). In 2007, a slight increase (p-value = 0.066) in resistance to GLS was detected in P30 C4 compared with C0, and no difference in resistance in P29. This same trend for GLS resistance was not apparent in 2008 (**Figures 6.3a and b**). It is notable that P30 C0 was more susceptible for both SLB and GLS than either cycle from P29. This could account for some of the increase in resistance in P30 as cycle 4 from P30 had the same level of resistance as P29.

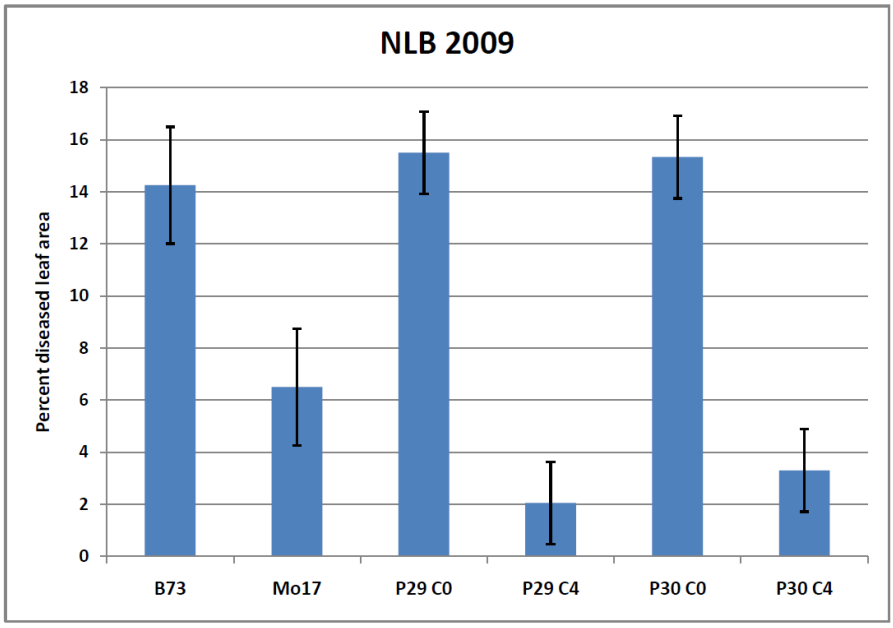


Figure 6.1. Confirmation of improved NLB resistance

Evaluation of cycle 0 and cycle 4 from P29 and P30 populations *per se* evaluated for NLB resistance. Inbred controls B73 and Mo17 are included.

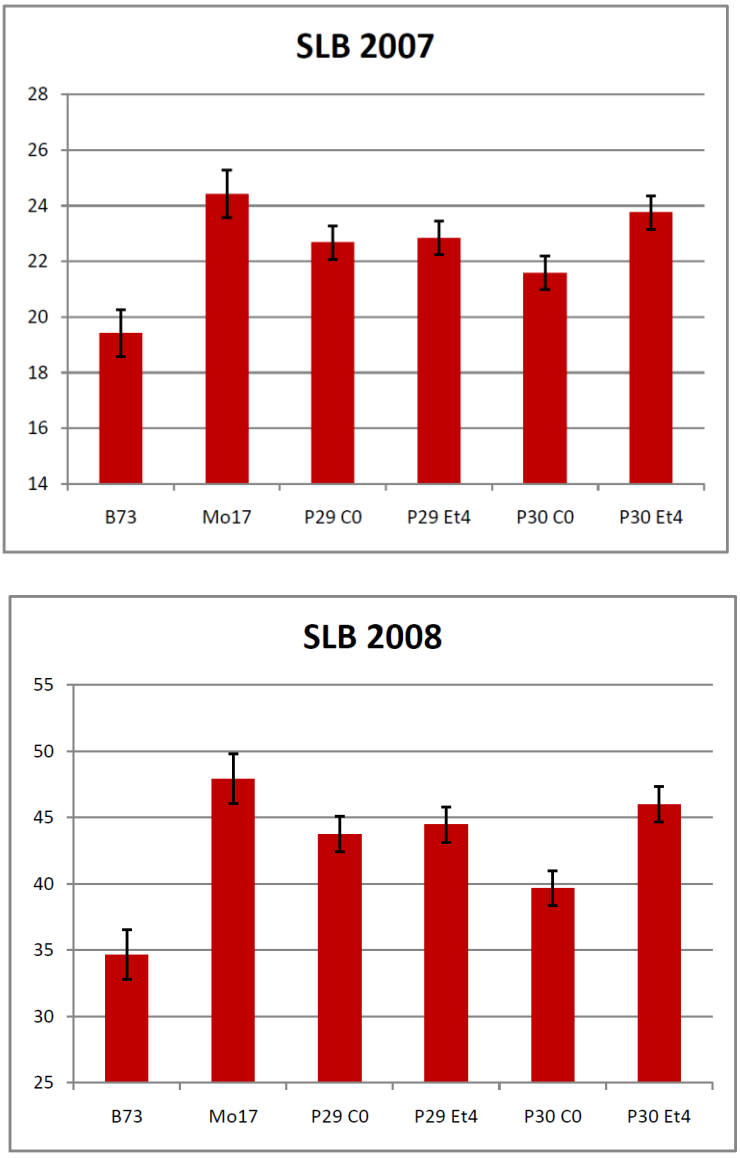
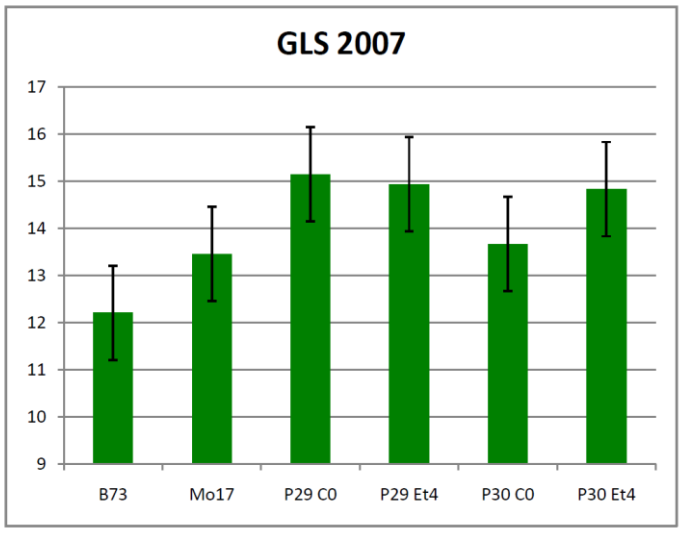


Figure 6.2. Improvement for SLB resistance in P30

Evaluation of P29 and P30 cycle 0 and cycle 4 for resistance to SLB in 2007 and 2008. Inbred controls B73 and Mo17 are included. SLB was an unselected disease in the original population improvement. The left axis shows the sum of all disease ratings. Higher values indicate greater resistance.



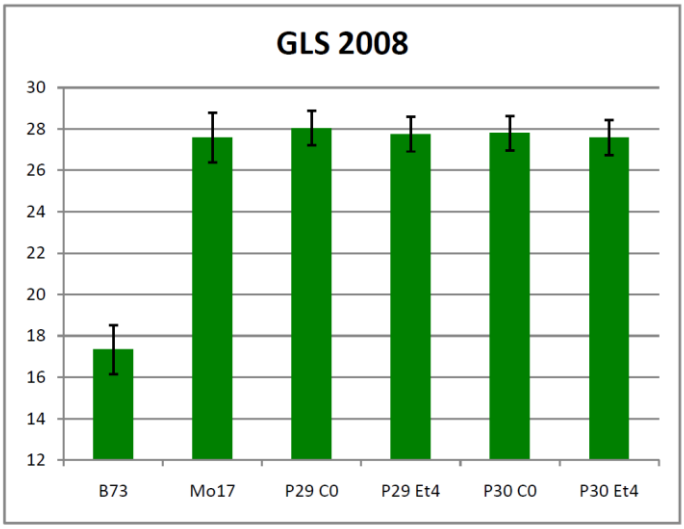


Figure 6.3. Evaluation for GLS resistance in recurrent selection populations.

Evaluation of P29 and P30 cycle 0 and cycle 4 for resistance to GLS in 2007 (a) and 2008 (b). Inbred controls B73 and Mo17 are included. GLS was an unselected disease in the original population improvement. The left axis shows area under the disease progress curve. Higher values indicate greater resistance.

Comparison to NAM. The 2009 NLB resistance trials for the two RS-RIL populations were planted alongside a larger nested association mapping (NAM) trial (Chap.2). The two trials were planted and inoculated on the same days. This set-up provided a direct comparison of the RS-RIL populations with the phenotypic diversity in NAM. For the third disease rating, the two RS-RIL populations had mean ratings (% diseased leaf area) of 21.0 and 16.4 respectively. At the same time-point, NAM had an average rating of 29.8 (**Figure 6.4**). Some of the more resistant NAM families such as CML277 and NC358 had lower average ratings of 11.7 and 17.9 respectively. However, this comparison shows that the RS-RIL populations fall on the resistant side of the NAM spectrum. Interestingly, the RS-RIL populations were very early maturing compared to NAM (**Figure 6.5**). The average number of days to anthesis (DTA) for Pop5 and Pop10 were 79.1 and 77.6 respectively. For NAM, the average DTA was 89.7, almost two weeks later than Pop10. Further, the two RS-RIL populations were earlier flowering than the earliest NAM family, P39 x B73, which averaged 84.0 days in the 2009 trial.

Correlation of disease resistances and relative maturity. The BLUPs for resistance to NLB, SLB and GLS for the two populations were compared for possible correlations. Correlated phenotypic response in a segregating RIL population is indication of either linked genes for the two traits or pleiotropic genes. There were significant positive correlations between resistance levels for each pari of diseases (**Figures 6.6a and b**). The correlations were in the range of 0.21 to 0.31 and were all significant at $p < 0.01$. Relative maturity (DTA) was also compared to disease resistance. In these pathosystems, resistance is generally negatively correlated with relative maturity (i.e. early maturity = more disease).

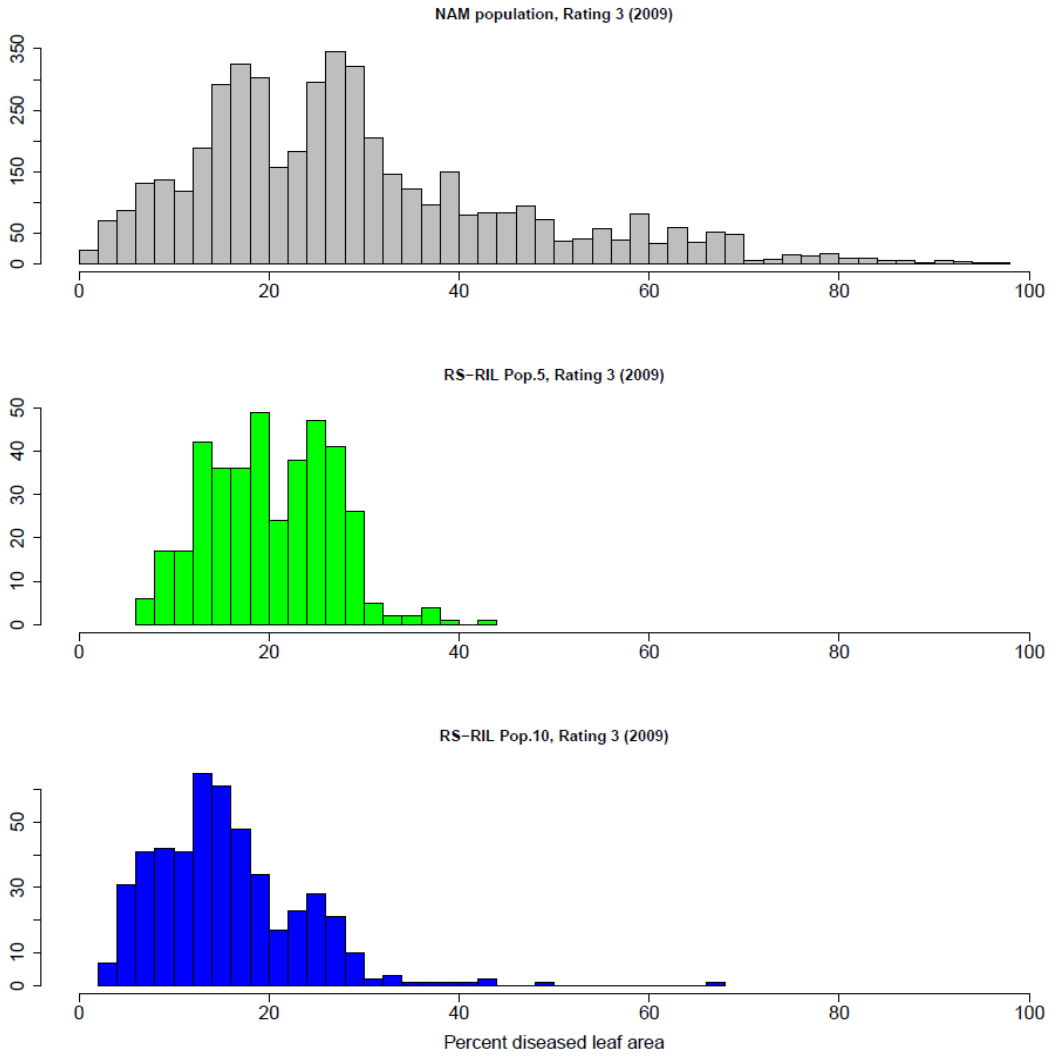


Figure 6.4. Comparison of NLB resistance in NAM and RS-RIL populations

Trait distribution for NAM, Pop5, and Pop10 evaluated for NLB resistance under the same experimental conditions in 2009. Both RS-RIL populations tend toward the resistant end of the NAM distribution. Pop10 is more resistant than Pop5. Several individual NAM families are more resistant than the RS-RIL populations. The number of lines observed at each rating level is shown on the left.

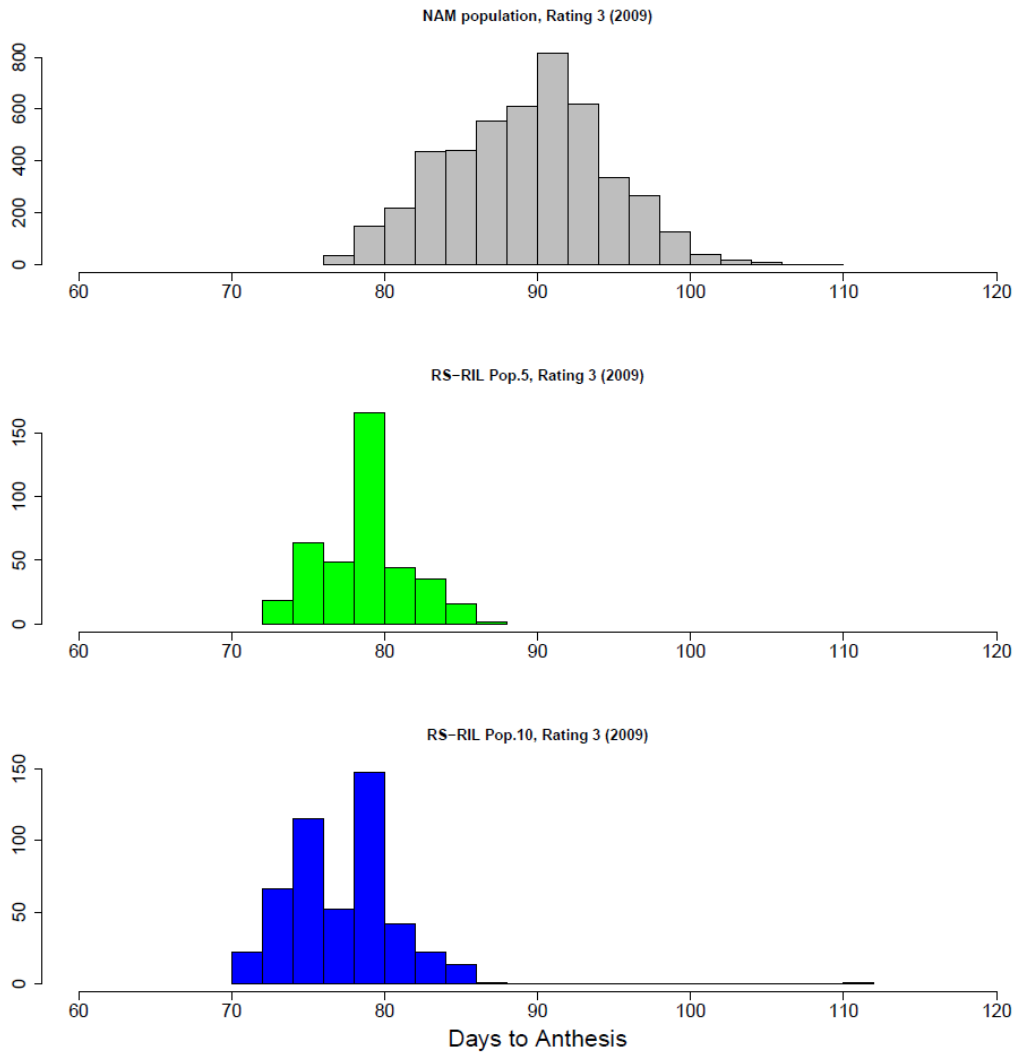
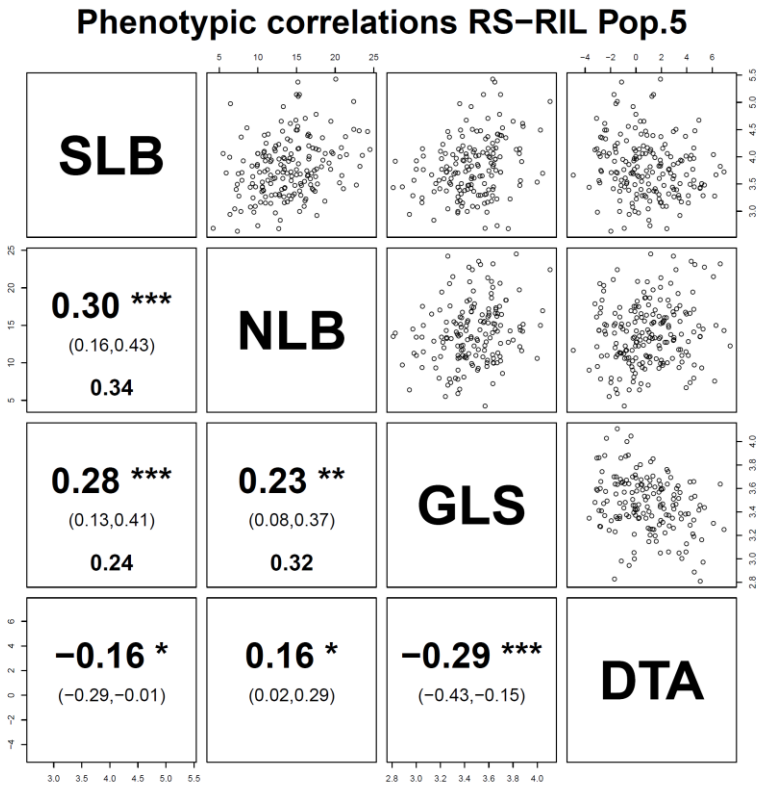


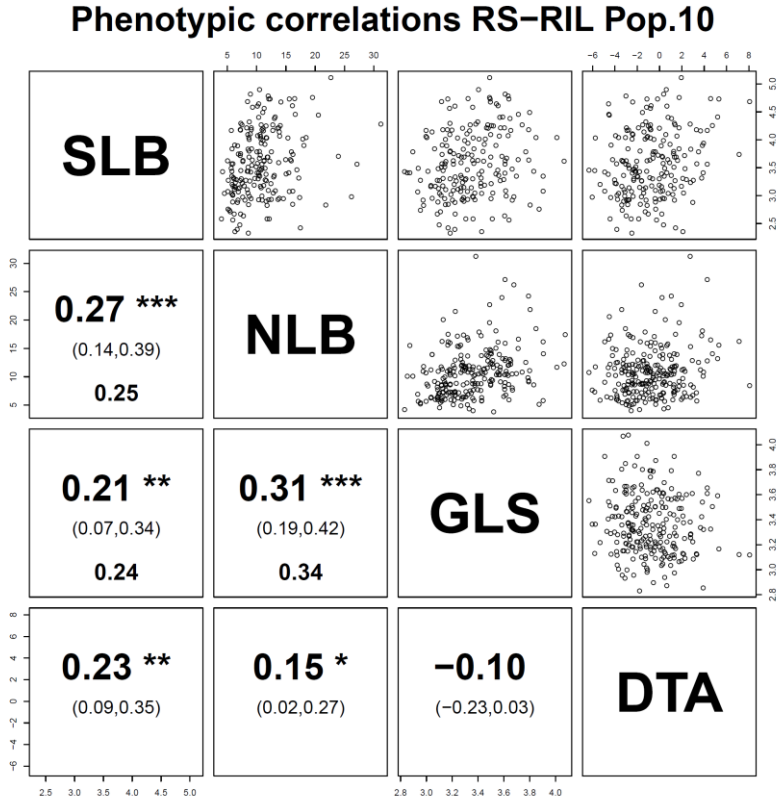
Figure 6.5. Comparison of relative maturity in NAM and RS-RIL populations

Trait distribution of days to anthesis (DTA) for NAM, Pop5, and Pop10. DTA was evaluated in a trial for NLB resistance under the same experimental conditions in 2009. Both RS-RIL populations are earlier than the earliest NAM family. The number of lines observed at each rating level is shown on the left.

Figure 6.6 Correlation between resistances for NLB, SLB and GLS in RS-RILs

Phenotypic correlations between resistances for NLB, SLB, and GLS in the two RIL populations. Correlation with relative maturity (days to anthesis; DTA) is also shown. 95% confidence interval for the estimated correlation is shown in brackets below the correlation value. Partial correlations accounting for DTA are shown at the bottom for disease – disease correlations.


Phenotypic correlations RS−RIL Pop.5
SLB
0.30 ***
(0.16,0.43)
0.34
NLB
0.28 ***
(0.13,0.41)
0.24
0.23 **
(0.08,0.37)
0.32
GLS
−0.16 *
(−0.29,−0.01)
0.16 *
(0.02,0.29)
−0.29 ***
(−0.43,−0.15)
DTA


Phenotypic correlations RS−RIL Pop.10
SLB
0.27 ***
(0.14,0.39)
0.25
NLB
0.21 **
(0.07,0.34)
0.24
0.31 ***
(0.19,0.42)
0.34
GLS
0.23 **
(0.09,0.35)
0.15 *
(0.02,0.27)
−0.10
(−0.23,0.03)
DTA

Significant negative correlations between DTA and disease resistance were found in Pop5 for SLB (-0.16) and GLS (-0.29), consistent with previous observations in these pathosystems. Surprisingly, NLB resistance was positively correlated with DTA in both Pop5 and Pop10, and in Pop10 SLB resistance was positively correlated with DTA.

DISCUSSION

To validate previous results from selection mapping for quantitative disease resistance in maize, two recombinant inbred populations were developed from recurrent selection material (RS-RILs). These two populations were derived from a single highly heterozygous individual from the last of four cycles of full-sib S_1 recurrent selection. From two different F_1 individuals, unique RIL populations were developed. These two populations represent an immortal genetic resource for dissection of resistance selected in the RS population. Further resources of >300 BC_2 lines were concurrently developed to allow genetic isolation and characterization of QRL identified in the RS-RIL populations.

Evaluation of the RS population *per se* confirmed the increased resistance to NLB from four cycles of recurrent selection. To examine the possibility of multiple disease resistance (MDR), the populations *per se* were evaluated for resistance to SLB and GLS, two diseases that were not selected during the original population improvement. Two years of evaluation confirmed increased resistance to SLB in P30 cycle 4 relative to P30 cycle 0. The same trend was not seen in P29. For GLS resistance in P30, a slightly higher level of resistance in C_4 relative to C_0 was observed in the 2007 trials but this was not confirmed in 2008. There was not any apparent increase in

resistance to GLS in P29. Concurrent improvement of an unselected trait indicates a positive genetic correlation in P30 for NLB and SLB. There was also limited evidence of the same correlation in P30 for NLB and GLS.

During the summer of 2009, the RS-RIL populations were evaluated for resistance to NLB, SLB and GLS. DTA measurements were also taken as a covariate to control for maturity-related effects on resistance. The NLB trial was grown along side of NAM, allowing comparison of the phenotypic diversity between NAM and the RS-RILs. NAM consistent of 25 RIL families from 25 diverse inbred parents. As a result NAM captures large phenotypic variation for NLB resistance. The RS-RIL populations have B73 as one parent, which also allows for reasonable comparison to NAM where the common reference parent is B73. The two RS-RIL populations fell toward the resistant end of the NAM spectrum, though several of the most resistant NAM populations had phenotypic distributions lower than the RS-RILs. However, the notable difference between the RS-RILs and NAM was the very early maturity of these populations. Pop10, for example, was almost two weeks earlier on average than NAM and almost one week earlier than the earliest NAM population, P39 xB73, in the 2009 trial. The RS-RIL populations are S_4 material while the NAM is S_5 material. The additional inbreeding of NAM could account for some of the later maturity. The early maturity of the RS-RIL populations is notable because early maturity is generally linked with increased susceptibility. In NAM, the early families (e.g. P39, IL14H, Oh43, MS71) are also more susceptible (see Chpt.2). The RS-RILs are not the most resistant material when compared to NAM as several NAM families have a lower average disease severity and distribution. However, when relative maturity is taken into account, the RS-RIL populations appear to constitute a unique slice of

maize diversity for disease resistance.

To examine the possibility of MDR loci in the RS-RIL populations, phenotypic correlations between the different disease resistances were examined. There were significant positive correlations between resistance for each of the diseases, supporting the idea that either linkage or pleiotropy contributes to the resistance phenotype. The phenotypic correlations between relative maturity and resistance were also evaluated for each of the diseases. Negative correlations between flowering time and GLS and SLB for Pop5 were consistent with previous work in these pathosystems. Surprisingly, there was a positive correlation between NLB resistance (both Pop5 and Pop10) and SLB resistance (Pop10) and DTA. This is in contrast to the typical negative association of disease resistance and relative maturity. During the original RS population improvement, there was a negative change in relative maturity while improving disease resistance (Ceballos *et al.* 1991). This RS population and the RS-RIL populations developed here indicate that the negative association between relative maturity and disease resistance can be broken.

The RIL populations (and complementary BC lines) developed here represent a useful genetic resource for further study of quantitative disease resistance in maize for NLB as well as SLB, GLS and other diseases. Concurrent studies indicate that these RS-RIL populations represent novel genetic material relative to the robust community maize resource NAM. Further phenotypic evaluation of these populations combined with genomic marker data will permit identification of resistance alleles and potentially alleles for MDR.

REFERENCES

Ceballos, H., J. A. Deutsch and H. Gutierrez (1991). Recurrent selection for resistance to *Exserohilum turcicum* in eight subtropical maize populations. *Crop Science* **31**: 964-971.

Coque, M. and A. Gallais (2006). Genomic regions involved in response to grain yield selection at high and low nitrogen fertilization in maize. *TAG Theoretical and Applied Genetics* **112** (7): 1205-1220.

Wisser, R. J., S. C. Murray, J. M. Kolkman, H. Ceballos and R. J. Nelson (2008). Selection Mapping of Loci for Quantitative Disease Resistance in a Diverse Maize Population. *Genetics* **180** (1): 583-599.